N° 89-654-X au catalogue — N° 2014001
ISBN 978-0-660-21389-7

Concepts et méthodes

Enquête canadienne sur l'incapacité, 2012 : Guide des concepts et méthodes

Division de la statistique sociale et autochtone

Février 2014

Statistique Canada Statistics Canada

Canada

Comment obtenir d'autres renseignements

Pour toute demande de renseignements au sujet de ce produit ou sur l'ensemble des données et des services de Statistique Canada, visiter notre site Web à www.statcan.gc.ca.

Vous pouvez également communiquer avec nous par :

Courriel à infostats@statcan.gc.ca

Téléphone entre 8 h 30 et 16 h 30 du lundi au vendredi aux numéros sans frais suivants :

- Service de renseignements statistiques 1-800-263-1136
- Service national d'appareils de télécommunications pour les malentendants 1-800-363-7629
- Télécopieur 1-877-287-4369

Programme des services de dépôt

- Service de renseignements 1-800-635-7943
- Télécopieur 1-800-565-7757

Comment accéder à ce produit

Le produit nº 89-654-X au catalogue est disponible gratuitement sous format électronique. Pour obtenir un exemplaire, il suffit de visiter notre site Web à www.statcan.gc.ca et de parcourir par « Ressource clé » > « Publications ».

Normes de service à la clientèle

Statistique Canada s'engage à fournir à ses clients des services rapides, fiables et courtois. À cet égard, notre organisme s'est doté de normes de service à la clientèle que les employés observent. Pour obtenir une copie de ces normes de service, veuillez communiquer avec Statistique Canada au numéro sans frais 1-800-263-1136. Les normes de service sont aussi publiées sur le site www.statcan.gc.ca sous « À propos de nous » > « Notre organisme » > « Offrir des services aux Canadiens ».

Publication autorisée par le ministre responsable de Statistique Canada

© Ministre de l'Industrie, 2014

Tous droits réservés. L'utilisation de la présente publication est assujettie aux modalités de l'entente de licence ouverte de Statistique Canada (http://www.statcan.gc.ca/reference/copyright-droit-auteur-fra.htm).

This publication is also available in English.

Note de reconnaissance

Le succès du système statistique du Canada repose sur un partenariat bien établi entre Statistique Canada et la population du Canada, ses entreprises, ses administrations et les autres établissements. Sans cette collaboration et cette bonne volonté, il serait impossible de produire des statistiques exactes et actuelles.

Signes conventionnels

Les signes conventionnels suivants sont employés dans les publications de Statistique Canada :

.	indisponible pour toute période de référence
..	indisponible pour une période de référence précise
...	n'ayant pas lieu de figurer
0	zéro absolu ou valeur arrondie à zéro
0[s]	valeur arrondie à 0 (zéro) là où il y a une distinction importante entre le zéro absolu et la valeur arrondie
[p]	provisoire
[r]	révisé
x	confidentiel en vertu des dispositions de la *Loi sur la statistique*
[E]	à utiliser avec prudence
F	trop peu fiable pour être publié
*	valeur significativement différente de l'estimation pour la catégorie de référence ($p < 0,05$)

Table des matières

1. Introduction **5**

1.1 Aperçu de l'enquête 5
1.2 Contexte 5
1.3 Mise en garde à l'intention des utilisateurs 6
1.4 Objet du Guide des concepts et des méthodes 7

2. Contenu de l'enquête : concepts et questions **8**

2.1 Identification des personnes ayant une incapacité 8
2.2 Calcul des taux d'incapacité au Canada 10
2.3 Évaluation de la sévérité des incapacités 11
2.4 Portrait des adultes canadiens ayant une incapacité 12
2.5 Modules du questionnaire 13
2.6 Contenu couplé de l'Enquête nationale auprès des ménages de 2011 14
2.7 Régions géographiques 16

3. Plan d'échantillonnage de l'ECI **17**

3.1 Population cible et couverture 17
3.2 Période de référence 18
3.3 Base de sondage tirée de l'ENM 18
3.4 Plan d'échantillonnage 18
3.5 Domaines d'estimation et stratification 19
3.6 Répartition de l'échantillon 20
3.7 Tailles d'échantillons 21
3.8 Échantillon des NON 21

4. Collecte des données **23**

4.1 Période de référence 23
4.2 Communications 23
4.3 Mode de collecte 23
4.4 Supervision et contrôle de la qualité 24
4.5 Interviews par procuration 24
4.6 Problèmes particuliers 25
4.7 Taux de réponse 25

5. Traitement des données **27**

5.1 Saisie des données 27
5.2 Étapes du traitement 27
5.3 Nettoyage des enregistrements : enregistrements dans le champ de l'enquête et enregistrements complets 28
5.4 Recodage des variables et questions à réponses multiples 28

5.5 Enchaînement des questions : parcours des réponses, enchaînements valides et la non-réponse aux questions 30
5.6 Codage 31
5.7 Vérifications et imputations 31
5.8 Variables dérivées et variables externes couplées de l'ENM 32

6. Pondération et création des fichiers de données définitives 34

6.1 Pondération 34
6.2 Structure et contenu des fichiers 41
6.3 Ensembles de données finaux et dictionnaires de données 42
6.4 Lignes directrices pour l'analyse 43

7. Qualité des données 44

7.1 Aperçu de l'évaluation de la qualité des données 44
7.2 Erreurs d'échantillonnage et méthode Bootstrap 44
7.3 Erreurs non dues à l'échantillonnage 46

8. Différences entre l'Enquête canadienne sur l'incapacité (ECI) de 2012 et l'Enquête sur la participation et les limitations d'activités (EPLA) de 2006 49

8.1 Nouvelle méthode d'identification des incapacités 49
8.2 Autres changements touchant le contenu de l'ECI 50
8.3 Questions filtres de l'ENM 51
8.4 Changement relatif au délai suite à l'enquête de présélection 51
8.5 Autres changements d'ordre méthodologique 51
8.6 Résumé et recommandation 52

9. Diffusion des données 53

9.1 Produits et services de données 53
9.2 Produits de référence 53
9.3 Contrôle de la divulgation 54

Annexe A – Indicateurs du sondage selon le module du questionnaire 55

Annexe B – Indicateur de sévérité des incapacités 61

Annexe C – Questions sur l'équité en matière d'emploi 66

Annexe D – Catégories supplémentaires de codage et classifications types utilisées 67

Annexe E – Glossaire des termes de l'enquête 69

Références 78

1. Introduction

1.1 Aperçu de l'enquête

L'Enquête canadienne sur l'incapacité (ECI) de 2012 est une enquête nationale qui vise les adultes canadiens dont les activités quotidiennes sont limitées par un état ou un problème de santé de longue durée. Elle a été élaborée par Statistique Canada avec le concours d'Emploi et Développement social Canada (EDSC) (anciennement Ressources humaines et Développement des compétences Canada). Les renseignements ayant servi à la conception de l'enquête ont été recueillis auprès du Groupe consultatif technique sur les personnes ayant une incapacité d'EDSC, composé d'experts du domaine de l'incapacité, notamment des universitaires et des représentants de diverses associations communautaires des différentes régions du Canada, de même que d'employés d'EDSC et de Statistique Canada. L'enquête a été menée du 24 septembre 2012 au 13 janvier 2013.

L'ECI repose sur un modèle social de l'incapacité plutôt que sur un modèle médical. Le modèle social est fondé sur le principe selon lequel l'incapacité résulte de l'interaction entre les limitations fonctionnelles d'une personne et les obstacles auxquels elle fait face dans l'environnement, notamment les obstacles sociaux et physiques qui compliquent la vie quotidienne. Par conséquent, l'incapacité est un désavantage social imposé par un environnement défavorable qui s'ajoute à la limitation fonctionnelle d'une personne (Mackenzie et coll., 2009).

L'ECI de 2012 offre un éventail de données sur différents types d'incapacité et leur sévérité. Elle évalue à quelle fréquence les activités quotidiennes des adultes canadiens sont limitées par des problèmes chroniques, des problèmes de santé et des difficultés liées à certaines tâches. L'enquête recueille également des données sur l'utilisation d'aides et d'appareils fonctionnels ainsi que l'aide reçue ou requise par les répondants. Elle comprend des questions sur la scolarité et l'expérience de travail des personnes ayant une incapacité, les aménagements faits à cet égard et la capacité à se déplacer dans la communauté.

Les données de l'ECI seront utilisées par les analystes des politiques relatives aux personnes ayant une incapacité et des politiques sociales de tous les ordres de gouvernement, ainsi que par les associations de personnes ayant une incapacité et les chercheurs qui œuvrent dans le domaine des politiques et des programmes touchant les personnes ayant une incapacité. Elles serviront à planifier et à évaluer les politiques, les programmes et les services à l'intention des adultes canadiens ayant une incapacité, en vue de les aider à participer pleinement à la vie en société. Notamment, l'information sur les adultes ayant une incapacité est essentielle à l'élaboration efficace et au bon fonctionnement du Programme d'équité en matière d'emploi. Les données sur l'incapacité servent également à respecter l'accord international du Canada lié à la Convention des Nations Unies relative aux droits des personnes ayant une incapacité .

L'Enquête canadienne sur l'incapacité (ECI) de 2012 a été menée à partir d'un échantillon de personnes qui ont déclaré une limitation d'activités dans l'Enquête nationale auprès des ménages (ENM) de 2011 et qui étaient âgées de 15 ans et plus à la date de l'ENM, soit le 10 mai 2011. Étant donné que l'ENM exclut la population vivant en établissement et les autres logements collectifs, l'ENM ne porte que sur les personnes qui habitent dans des logements privés au Canada. De plus, pour des raisons opérationnelles, la population des réserves des Premières Nations est également exclue. La taille totale de l'échantillon de l'ECI était de quelque 45 500 personnes. Le taux de participation global a atteint 75 %. L'ECI fournit des données fiables sur les personnes ayant une incapacité dans chaque province et territoire au Canada. Les groupes d'âge sont également ventilés en fonction de la province.

1.2 Contexte

Les données sur les limitations d'activités au Canada sont recueillies depuis une trentaine d'années, ce qui traduit bien la reconnaissance grandissante de l'importance des données pour soutenir l'objectif fixé, aider les personnes ayant une incapacité à participer pleinement à la vie en société. Les concepts et les méthodes employés dans la production de données sur les incapacités ont évolué avec le temps. La première enquête sur l'incapacité au Canada a été menée en 1981, l'Année internationale des personnes handicapées, peu après la publication du rapport intitulé « Obstacles » par le Comité spécial du Parlement concernant les invalides et les handicapés. Les auteurs du rapport ont recommandé, entre autres choses, que Statistique

Canada produise des données sur les personnes ayant une incapacité. C'est ce rapport qui a débouché sur la mise en place de l'Enquête sur la santé et l'incapacité, qui se voulait un complément à l'Enquête sur la population active d'octobre 1983 et de juin 1984 . De plus, le Recensement de 1986 contenait une question concernant les limitations d'activités, qui devait aider à identifier les personnes susceptibles d'avoir une incapacité. Un peu plus tard la même année, Statistique Canada s'est servi de la base du recensement afin de constituer un échantillon pour l'Enquête sur la santé et les limitations d'activités (ESLA). Cette enquête a été conçue de manière à identifier les Canadiens ayant une incapacité et à déterminer leurs limitations et les obstacles auxquels ils faisaient face. Une deuxième ESLA a eu lieu en 1991.

En 1996, aucune enquête postcensitaire n'a été effectuée. En 1998, les gouvernements fédéral, provinciaux et territoriaux ont toutefois dévoilé un cadre stratégique commun pour les personnes ayant une incapacité, intitulé *À l'unisson*, qui favorisait une plus grande inclusion des personnes ayant une incapacité dans tous les aspects de la société canadienne. Leur rapport de 1998 soulignait l'importance de l'élaboration d'une base de données statistiques fiables et le rôle essentiel que les données d'enquête jouerait dans le soutien à l'élaboration de politiques et à la recherche dans ce domaine.

En 2001, la Classification internationale du fonctionnement, du handicap et de la santé (CIF) a été approuvée par les états membres de l'Organisation mondiale de la santé (OMS), y compris le Canada. La CIF définit l'incapacité en rapport avec la relation qui s'établit entre les structures et les fonctions corporelles, les activités de la vie quotidienne et la participation sociale tout en tenant compte du rôle des facteurs du milieu. Reflet de cette nouvelle définition de l'incapacité, axée sur une approche liée à un modèle social, l'ESLA a été rebaptisée Enquête sur la participation et les limitations d'activités (EPLA) et elle a été menée en 2001 et en 2006. Comme l'indique son nouveau nom, l'enquête mettait l'accent sur la participation des personnes dont les activités sont limitées. À l'instar de l'ESLA, l'EPLA est un projet conjoint d'Emploi et Développement social Canada et de Statistique Canada.

Tout récemment, en 2010, le Canada a ratifié la Convention des Nations Unies relative aux droits des personnes ayant une incapacité. Conformément à l'Article 31 sur les statistiques et la collecte des données, EDSC a lancé la nouvelle stratégie en matière de données pour les personnes ayant une incapacité et a élaboré une série de questions afin d'identifier les personnes ayant une incapacité, ce qu'on a appelé les « Questions d'identification des incapacités » (QII). Les QII sont plus axées sur le modèle social de l'incapacité afin d'uniformiser davantage les catégories d'incapacité et d'améliorer la couverture de la gamme complète des catégories d'incapacité, notamment les incapacités mentales ou psychologiques et les troubles d'apprentissage et de la mémoire. Les QII ont été l'objet d'essais qualitatifs et quantitatifs poussés et elles ont ensuite été utilisées pour la première fois en vue d'identifier les adultes ayant une incapacité dans l'Enquête canadienne sur l'incapacité (ECI) de 2012. L'ECI a également brossé un tableau des adultes canadiens ayant une incapacité par rapport à leur participation à la vie en société.

1.3 Mise en garde à l'intention des utilisateurs

Les concepts et les méthodes employés pour l'évaluation de l'incapacité dans l'Enquête canadienne sur l'incapacité (ECI) de 2012 se démarquent considérablement de ceux utilisés dans l'Enquête sur la participation et les limitations d'activités (EPLA) de 2006. Le changement le plus important réside dans le fait que la définition de l'incapacité n'est pas la même dans les deux enquêtes. Dans le cadre de l'ECI, la nouvelle définition s'appliquait au moyen de la nouvelle série de questions d'identification des incapacités, traduisant une mise en oeuvre plus complète du modèle social de l'incapacité. D'autres changements majeurs apportés à l'ECI (se reporter au chapitre 8 ci-dessous) se répercutent également sur la comparabilité des séries chronologiques. Compte tenu des grandes différences entre les concepts et les méthodes employés dans l'EPLA de 2006 et l'ECI de 2012, il n'est ni possible ni recommandé de comparer la prévalence de l'incapacité au fil du temps entre ces deux sources de données.

1.4 Objet du Guide des concepts et des méthodes

Le Guide des concepts et des méthodes est conçu pour aider à comprendre l'Enquête canadienne sur l'incapacité (ECI) de 2012 en ce qui concerne son sujet et ses approches méthodologiques. Il est conçu pour aider les utilisateurs de données de l'ECI en servant de guide en ce qui a trait aux concepts et questions employés dans l'enquête ainsi qu'aux détails techniques de la conception de l'enquête, du travail de terrain et du traitement des données. Ce guide contient des renseignements utiles quant à la façon d'utiliser et d'interpréter les résultats d'enquête. La description de la qualité des données permet également aux utilisateurs d'examiner les points forts et les limites des données par rapport à leurs propres besoins.

Le chapitre 1 de ce guide donne un aperçu de l'ECI de 2012 en présentant le contexte et les objectifs de l'enquête. Le chapitre 2 traite de l'élaboration du contenu de l'enquête et des essais en expliquant les concepts clés et les définitions employés dans l'enquête. Ce chapitre présente les modules du questionnaire de l'ECI ainsi que les couplages d'informations avec l'Enquête nationale auprès des ménages (ENM). Les chapitres 3 à 6 traitent d'aspects importants de la méthodologie de l'enquête, du plan de sondage à la création des fichiers de données définitives, en passant par la collecte et le traitement des données. Les chapitres 7 et 8 traitent des problèmes liés à la qualité des données et ils présentent des mises en garde aux utilisateurs quant aux comparaisons avec des données des enquêtes sur la participation et les limitations d'activités (EPLA) antérieures. Le chapitre 9 décrit les résultats de l'enquête mis à la disposition du grand public, notamment les tableaux de données, un feuillet d'information et des documents de référence. Les annexes offrent des précisions sur les indicateurs du questionnaire, l'indicateur de sévérité des incapacités, les questions utilisées afin d'évaluer l'équité en matière d'emploi, les catégories de codage spécial pour l'enquête et les classifications types utilisées. Ce document contient également un glossaire des termes utilisés dans l'enquête.

2. Contenu de l'enquête : concepts et questions

2.1 Identification des personnes ayant une incapacité

Dans la foulée de la nouvelle stratégie en matière de données sur les personnes ayant une incapacité de 2010, on a créé les « Questions d'identification des incapacités» (QII), un outil visant à mieux identifier les personnes ayant une incapacité. Les QII représentent la première composante de la version plus complète de l'Enquête canadienne sur l'incapacité.

En élaborant ces nouvelles questions d'identification, on visait à mettre au point un nouveau module d'identification complet des personnes ayant une incapacité, présentant les caractéristiques suivantes :

- être orienté davantage vers le modèle social de l'incapacité;
- déterminer le type et la sévérité de l'incapacité;
- uniformiser davantage les catégories d'incapacité;
- améliorer la couverture de toute la gamme des types d'incapacité, notamment les incapacités mentales ou psychologiques ainsi que les troubles d'apprentissage et de mémoire;
- être suffisamment court pour s'adapter à toute une gamme d'enquêtes à l'échelle de l'ensemble de la population.

Il convient de souligner que les QII ont été précisément mises au point pour évaluer les incapacités chez les adultes canadiens, mais non chez les enfants.

Questions d'identification des incapacités (QII)

Le modèle social de l'incapacité indique qu'à elle seule, une difficulté ne suffit pas à définir une incapacité; une limitation dans la vie quotidienne doit également être déclarée. Par conséquent, les QII évaluent le type et la sévérité des incapacités d'adultes canadiens à qui l'on demande de préciser à quelle fréquence leurs activités quotidiennes sont limitées par une condition physique ou mentale de longue durée, des problèmes de santé et des difficultés dans l'accomplissement de tâches.

Les questions d'identification dans les QII permettent d'évaluer la présence et la sévérité de dix catégories distinctes d'incapacités liées à des difficultés ou à des problèmes de santé qui durent depuis au moins six mois ou qui pourraient durer six mois ou plus. Les questions d'identification mettent l'accent sur l'uniformité des mesures entre les catégories d'incapacité. Ces questions portent sur les catégories d'incapacité suivantes :

1. Vision
2. Ouïe
3. Mobilité
4. Flexibilité
5. Dextérité
6. Douleur
7. Apprentissage
8. Développement
9. Mental/psychologique
10. Mémoire

On retrouve également dans les QII une question au sujet de tout autre problème de santé ou condition qui dure ou pourrait durer six mois ou plus. On associe cette question au type « Autre ». Toutefois, on tiendra compte de ce type seulement si aucune autre limitation n'a été déclarée aux 10 types d'incapacité listés ci-dessus. Si au contraire il existe une limitation à un des 10 types et une limitation au type « Autre », alors ce dernier sera ignoré[1]. On a donc créé un type dit « Inconnu » pour les personnes ayant déclaré une limitation de type « Autre » seulement et aucune autre limitation.

Mise à l'essai sur les plans qualitatif et quantitatif des questions d'identification des incapacités

On a rédigé les nouvelles questions d'identification des incapacités après avoir étudié en profondeur les indicateurs d'incapacité actuellement utilisés au Canada etau niveau international. Ensuite, les QII ont été testées sur les plans qualitatif et quantitatif. Différentes itérations des essais qualitatifs ont été effectuées par le Centre de ressources en conception de questionnaires (CRCQ) de Statistique Canada en septembre 2010, en mars 2011 et en mai 2011. Les essais qualitatifs ont été effectués en français et en anglais, à six endroits différents au Canada, ce qui, dans l'ensemble, représente environ 125 interviews cognitifs individuels et quatre groupes de discussion. Des modifications ont été apportées à la formulation des questions en fonction de ces résultats.

Au cours de l'automne 2011, on a également soumis les QII à des essais quantitatifs dans deux enquêtes de Statistique Canada : l'Enquête sur la population active et l'Enquête sur la santé dans les collectivités canadiennes. Ces essais ont démontré que les QII étaient un instrument fiable pour estimer la prévalence de l'incapacité dans la population adulte en fonction du modèle social d'incapacité. Après les dernières modifications à la formulation des questions et aux directives destinées aux intervieweurs, l'utilisation des QII dans l'Enquête canadienne sur l'incapacité (ECI) a été autorisée.

Il convient de souligner qu'une version plus courte des QII a également été élaborée comme option minimale pour identifier les personnes ayant une incapacité dans les enquêtes visant la population générale à Statistique Canada. On pourra ainsi identifier de façon fiable les personnes ayant une incapacité dans les enquêtes nationales sur la population générale plus fréquemment, en plus de l'Enquête canadienne sur l'incapacité (ECI), plus spécialisée. On peut donc produire des données plus variées sur les personnes ayant une incapacité au Canada.

Définition opérationnelle de l'incapacité au moyen des QII

Pour chacun des 10 types d'incapacité, les QII comprennent toujours au moins une question sur le niveau de difficulté relié à ce type d'incapacité (Aucune difficulté, Un peu de difficulté, Beaucoup de difficulté, Incapable de le faire) et une question sur la fréquence des limitations d'activité (Jamais, Rarement, Parfois, Souvent ou Toujours). Pour identifier une incapacité à un type donné, il faut que la fréquence des limitations soit « Parfois », « Souvent » ou « Toujours », ou qu'elle soit « Rarement » mais combinée à un niveau de difficulté de « Beaucoup » ou « Incapable de le faire ».

1 La décision d'ignorer le type « Autre » lorsqu'on a déjà une limitation à un des 10 types d'incapacité vient du fait qu'on a pu remarquer que les répondants ayant une incapacité de l'un des 10 types avaient tendance à déclarer, dans le type « Autre », la condition médicale responsable de leur incapacité. On voulait donc éviter de compter en double les types d'incapacité.

Le tableau 2.1 ci-dessous résume la combinaison de réponses dans les QII qui servent généralement à identifier une incapacité. Cette classification s'applique à la majorité des catégories d'incapacité évaluées par les QII.

Tableau 2.1
Combinaisons de réponses aux Questions d'identification des incapacités utilisées pour identifier une incapacité

Dans quelle mesure avez-vous de la difficulté à...?	À quelle fréquence vos activités sont-elle limitées...?				
	Jamais	Rarement	Parfois	Souvent	Toujours
Aucune difficulté	Pas d'incapacité	Pas d'incapacité	Incapacité	Incapacité	Incapacité
Un peu de difficulté	Pas d'incapacité	Pas d'incapacité	Incapacité	Incapacité	Incapacité
Beaucoup de difficulté	Pas d'incapacité	Incapacité	Incapacité	Incapacité	Incapacité
Incapable	Pas d'incapacité	Incapacité	Incapacité	Incapacité	Incapacité

Source : Statistique Canada, Enquête canadienne sur l'incapacité, 2012.

Il convient de souligner que dans certains cas, ces critères ont été modifiés. Plus particulièrement, une personne qui déclare souffrir d'un trouble du développement est considérée comme ayant une incapacité si cet état a été diagnostiqué chez elle, peu importe le degré de difficulté ou la fréquence de la limitation d'activités déclarée.

Autre exception à souligner, l'incapacité de type « Inconnu », où nous ne posons pas de question sur le degré de difficulté. Ainsi une personne aura une incapacité de type « Inconnu » seulement si elle déclare être limitée « Parfois », « Souvent » ou « Toujours » à cause de ce problème de santé ou de cette condition et qu'elle ne déclare aucune autre limitation pour les 10 types d'incapacité.

Enfin, pour les incapacités concernant la vision, l'ouïe, la mobilité, la flexibilité et la dextérité, qui sont évaluées à l'aide de questions liées à des tâches, la réponse « Aucune difficulté » n'est pas suivie de la question sur les limitations relatives aux activités quotidiennes. Donc, toutes les réponses « Aucune difficulté » pour ces catégories d'incapacité sont classées dans la catégorie « Pas d'incapacité ».

2.2 Calcul des taux d'incapacité au Canada

L'un des principaux objectifs de l'ECI de 2012 consiste à calculer les taux d'incapacité au Canada. Ceux-ci peuvent se calculer en fonction de la province et du territoire, par exemple, ou du groupe d'âge. Voici la formule qui sert à les calculer :

(Personnes ayant une incapacité / (Personnes ayant une incapacité + Personnes sans incapacité)) x 100

Afin d'obtenir ce genre de statistiques, il fallait que la méthodologie de l'ECI réussisse non seulement à identifier les personnes ayant une incapacité, mais également à donner des estimations du nombre de personnes sans incapacité au Canada. Voilà pourquoi deux échantillons distincts de personnes ont été constitués à partir de l'Enquête nationale auprès des ménages (ENM) de 2011 :

1. Les personnes du premier échantillon (appelé l'échantillon OUI) ont été filtrées à l'aide des questions de l'ENM sur les limitations d'activités pour être ensuite soumises aux Questions d'identification des incapacités (QII) de l'Enquête canadienne sur l'incapacité (ECI), ce qui a permis d'identifier plus précisément les personnes ayant une incapacité.
2. Les personnes du deuxième échantillon (appelé l'échantillon NON) ont été exclues du premier groupe à l'aide des questions de l'ENM sur les limitations d'activités et étaient automatiquement considérées comme des personnes sans incapacité.

Ces méthodes sont décrites en détail au chapitre 3 du guide.

2.3 Évaluation de la sévérité des incapacités

Utilité de l'indice de sévérité

Selon des études antérieures faites à partir de l'Enquête sur la participation et les limitations d'activités de 2006, il est évident que la sévérité des incapacités est un excellent indicateur de la réduction des activités des personnes ayant une incapacité dans plusieurs domaines de la vie quotidienne (Rapport Fédéral sur les personnes handicapées – 2010). Il est beaucoup moins probable que les gens ayant une incapacité sévère soient actifs sur le marché du travail, qu'ils s'inscrivent à des programmes d'éducation postsecondaire et terminent leurs études et qu'ils participent à la vie communautaire, que leurs homologues souffrant d'une incapacité moins grave. Les gens ayant une incapacité sévère vont probablement avoir besoin de mesures de soutien et de services, comme des aides, des appareils fonctionnels, des soins, des médecins spécialistes et des mesures de soutien du revenu.

L'inclusion du degré de sévérité des incapacités représente un facteur important dans les analyses de la participation des personnes ayant une incapacité. Grâce à l'indice de sévérité des incapacités uniforme et facile à utiliser inclus dans le fichier de l'ECI, les analystes peuvent élaborer des inférences plus précises relativement à la situation à laquelle font face les personnes ayant une incapacité.

L'indice de sévérité

Un score de sévérité a été développé à partir des questions d'identification des incapacités (QII). À partir d'une grille de pointage qui tient compte à la fois de la fréquence des limitations d'activité (Jamais, Rarement, Parfois, Souvent ou Toujours) et de l'intensité des difficultés (Aucune difficulté, Un peu de difficulté, Beaucoup de difficulté ou Incapable de le faire), un score est assigné à chacun des 10 types d'incapacité. Ce score augmente avec la fréquence de la limitation et avec le niveau de difficulté éprouvé.

Un indice de sévérité global est dérivé à partir toutes les catégories d'incapacité. L'indice de sévérité global pour une personne s'obtient à partir de la moyenne des scores des 10 types d'incapacité. Ainsi plus le nombre de types d'incapacité est grand, plus son indice sera élevé.

Le score global développé se devait de répondre aux 3 critères suivants :

- le score doit augmenter avec le nombre de types d'incapacité
- le score doit augmenter avec le niveau de difficulté relié à l'incapacité
- le score doit augmenter avec la fréquence de la limitation des activités

Afin de faciliter l'utilisation du score de sévérité, des classes de sévérité ont été établies. Il est important de comprendre que le nom assigné à chaque classe a pour seul but d'en faciliter l'utilisation. Il ne constitue pas une forme d'étiquette ou de jugement du niveau d'incapacité de la personne. En d'autres termes, la façon d'interpréter ces classes est de dire que les personnes de la classe 1 ont une incapacité moins sévère que celles de la classe 2, ces dernières ayant une incapacité moins sévère que celles de la classe 3, et ainsi de suite.

- 1 = Incapacité légère
- 2 = Incapacité modérée
- 3 = Incapacité sévère
- 4 = Incapacité très sévère

Le tableau suivant donne la répartition des personnes avec incapacité selon les 4 classes de sévérité.

Tableau 2.2
Distribution des personnes avec incapacité selon la classe de sévérité, Canada

Classe de sévérité	Personnes avec incapacité	
	nombre	%
Classe 1 = léger	1 195 590	31,7
Classe 2 = modéré	747 980	19,8
Classe 3 = sévère	849 540	22,5
Classe 4 = très sévère	982 810	26,0

Source : Statistique Canada, Enquête canadienne sur l'incapacité, 2012.

Pour en savoir davantage sur les méthodes employées pour dériver les indices de sévérité, reportez-vous à l'annexe B.

2.4 Portrait des adultes canadiens ayant une incapacité

En plus des Questions d'identification des incapacités (QII), on a défini le contenu détaillé de l'Enquête canadienne sur l'incapacité (ECI) qui dressait le portrait des problèmes affectant les adultes canadiens qui ont de la difficulté à accomplir leurs activités quotidiennes. L'ECI portait sur l'utilisation d'aides et d'appareils fonctionnels ainsi que sur l'aide reçue ou requise par les personnes ayant une incapacité. Les priorités en matière de données comprenaient également l'expérience en matière d'éducation et d'emploi des adultes ayant une incapacité, les aménagements faits dans ces domaines, la capacité à se déplacer dans la collectivité en recourant aux transports en commun ainsi que les sources de revenus.

Les indicateurs de l'enquête dans ces domaines provenaient de sources variées, notamment l'Enquête sur la participation et les limitations d'activités (EPLA) de 2006. On a également recueilli des mesures normalisées et bien établies, employées dans d'autres enquêtes de Statistique Canada, qu'on a examinées en les considérant comme des indicateurs éventuels. Les indicateurs jugés particulièrement importants ont ensuite été modifiés pour qu'ils soient mieux adaptés à la population des personnes ayant une incapacité.

On a mis à l'essai un questionnaire provisoire en mars 2012. Les essais qualitatifs comportaient 50 interviews cognitives en français et en anglais, réalisées à trois endroits différents au Canada. Les résultats des essais ont débouché sur des recommandations concernant une série définitive de questions pour l'Enquête canadienne sur l'incapacité intégrale. On a apporté des améliorations à la formulation des questions, aux directives à l'intention des intervieweurs et à l'enchaînement des questions.

Dans le cadre de l'ECI de 2012, les questions ont été conçues de manière à être utilisées dans les interviews assistées par ordinateur (IAO). Les IAO permettent un enchaînement plus complexe des questions et intègrent de nombreuses caractéristiques visant à optimiser la qualité des données, comme des vérifications de la cohérence par comparaison des questions. Dans le cas de l'ECI, le questionnaire consistait en une interview téléphonique assistée par ordinateur (ITAO). La nouvelle application du questionnaire ITAO a été soumise à des essais rigoureux (essais modulaires et essais bout à bout).

Pour ce qui est de l'ECI de 2012, on a pris une dernière mesure concernant l'élaboration du contenu : couplage des données avec celles de l'Enquête nationale auprès des ménages (ENM) de 2011. Étant donné que l'échantillon de l'ECI de 2012 a été constitué à partir des données de l'ENM de 2011 (reportez-vous au chapitre 3 pour en savoir davantage), les données utiles provenant de l'ENM peuvent être associées aux données recueillies au cours de l'interview de l'ECI de 2012. Cette approche réduit le nombre de questions à poser dans l'ECI et permet de brosser un tableau précis des gens ayant une incapacité pour les utilisateurs de données de l'ECI. De plus, étant donné qu'on a constitué un échantillon de personnes sans incapacité à partir de l'ENM afin de calculer les taux d'incapacité au Canada (reportez-vous au chapitre 3 pour en savoir davantage), le couplage des données avec celles de l'ENM représente également une excellente source d'information sur les gens sans incapacité, ce qui permet d'établir des comparaisons importantes de caractéristiques entre les deux groupes.

Cette procédure de couplage des données de l'ENM est présentée plus en détail à la section 2.6 ci-dessous.

Remarque concernant les données sur l'âge

L'âge figure au nombre des principaux facteurs démographiques qui nous intéressent dans l'analyse des incapacités au Canada. Lorsqu'on utilise l'âge dans une étude effectuée avec les données de l'ECI de 2012 conjointement avec des données couplées provenant de l'ENM de 2011, il importe que les utilisateurs se rappellent les différentes périodes de référence visées. La section 6.2 du présent guide explique les périodes de référence de l'enquête. Pour ce qui est de l'âge, il convient de souligner que la collecte des données dans le cadre de l'ENM a été effectuée le 10 mai 2011, et celle de l'ECI, de 16 à 20 mois plus tard, soit entre septembre 2012 et janvier 2013. Ainsi, les répondants à l'ECI qui avaient 15 ans à l'époque de l'ENM avaient 16 à 20 mois de plus au moment de l'ECI. En ce qui a trait aux études particulières pour lesquelles l'écart entre les périodes de référence fait une différence, les utilisateurs de données auront la possibilité de sélectionner un indicateur d'âge établi en fonction de la période de référence de l'ENM de 2011 ou encore une variable d'âge en fonction de la date de l'interview de l'ECI de 2012. La section 6.2 explique l'utilisation des périodes de référence de l'enquête par rapport à différents types d'analyses de données susceptibles d'intéresser les utilisateurs.

2.5 Modules du questionnaire

Voici les modules du questionnaire qui se trouvent dans l'Enquête canadienne sur l'incapacité de 2012. La description détaillée de chacun d'eux figure à l'annexe A.

L'Enquête canadienne sur l'incapacité de 2012

Introduction

- Objet de l'enquête
- Nature volontaire de l'enquête
- Explication du couplage des données avec l'ENM

Questions d'identification des incapacités

- Questions filtres
- Questions de sélection

Questions sur les incapacités

- Principale condition sous-jacente
- Aides et appareils fonctionnels
- Questions récurrentes portant sur les aides et les appareils fonctionnels
- Consommation de médicaments
- Aide requise et reçue dans les activités quotidiennes

Études

- Fréquentation récente de l'école
- Fréquentation passée de l'école
- Expériences pédagogiques
- Antécédents scolaires

Emploi

- Situation par rapport au marché du travail
- Détails liés à l'emploi
- Personnes sans emploi
- Hors du marché du travail
- Retraite
- Formation en milieu de travail
- Modifications à l'emploi
- Discrimination en milieu de travail

Déplacements dans la ville ou communauté

Source de revenus

À l'information sur les modules du questionnaire, qui figure à l'annexe A, s'ajoutent les catégories de questions supplémentaires à l'annexe D, créées pendant le codage, de même que les classifications types qui ont servi à créer des indicateurs pour les questions ouvertes. L'Enquête Canadienne sur l'incapacité est accessible sur le site Web de Statistique Canada.

La description complète de toutes les variables dérivées des données de l'enquête se trouve dans les dictionnaires de données de l'ECI de 2012. Pour en savoir plus long sur la façon de se procurer les dictionnaires de données, communiquez avec les Services à la clientèle de Statistique Canada.

2.6 Contenu couplé de l'Enquête nationale auprès des ménages de 2011

L'échantillon de l'Enquête canadienne sur l'incapacité (ECI) a été constitué à partir de répondants à l'Enquête nationale auprès des ménages (ENM) de 2011 (reportez-vous au chapitre 3 pour en savoir davantage). Au début de l'interview de l'ECI de 2012, tous les répondants ont été mis au courant des projets de couplage des données de l'ECI avec l'information qu'ils avaient donné dans l'ENM. Toutes les données couplées demeureront confidentielles et ne serviront qu'à des fins statistiques.

Les avantages particuliers du couplage d'enregistrements ECI-ENM sont les suivants : diminution du fardeau de réponse pour la population cible de l'ECI, établissement des poids de l'enquête qui sont essentiels à la production d'estimations et création d'un fichier complet de microdonnées sur les personnes ayant une incapacité au Canada. Prises ensemble, les données de ces deux sources dressent un portrait statistique précis des personnes ayant une incapacité au Canada (des données qu'il est impossible d'obtenir d'autres sources).

Comme il est expliqué plus précisément au chapitre 3, on a également constitué un échantillon de personnes sans incapacité à partir de l'ENM pour l'ECI. Le couplage avec l'ENM permet donc aux utilisateurs de données de comparer certaines caractéristiques des personnes ayant une incapacité et celles qui n'en ont pas.

Quelque 200 variables de l'ENM ont été couplées avec la version finale du fichier de l'ECI pour 2012, pour les personnes ayant une incapacité et les personnes sans incapacité. Voici la liste des catégories de variables de l'ENM qui ont été couplées dans les fichiers analytiques de l'ECI.

- Variables au niveau du ménage
  - La géographie, y compris les régions métropolitaines de recensement
  - Le logement, notamment le mode d'occupation, le nombre de chambres dans le logement et les réparations nécessaires
  - La famille, notamment la présence d'un conjoint ou d'un partenaire, le nombre d'enfants dans la famille de recensement

- Variables au niveau de la personne
  - Les limitations d'activités, notamment les difficultés dans les activités ou la réduction des activités à la maison, au travail et à l'école
  - L'ascendance et l'identité autochtones
  - Les caractéristiques démographiques, notamment l'âge, l'état matrimonial et l'union libre, le lieu de naissance
  - L'immigration, la citoyenneté et l'origine ethnique
  - L'éducation, notamment la fréquentation scolaire, le lieu des études, les certificats postsecondaires, les grades et les diplômes, les types obtenus, le principal domaine d'études
  - L'emploi, notamment la situation par rapport au marché du travail, le nombre de semaines travaillées en 2010, la catégorie de travailleurs, le travail à plein temps ou à temps partiel
  - Le lieu de travail, notamment la catégorie de lieu de travail, le type de moyen de transport et la distance
  - La mobilité, notamment la mobilité un an et cinq ans auparavant
  - Le revenu, notamment le revenu familial, les revenus d'emploi, la situation de faible revenu (avant et après impôt)
  - La langue, notamment la connaissance des langues officielles, la langue maternelle, la langue parlée à la maison et la langue de travail.

Il convient de souligner que les variables de l'ENM renvoient à la situation de chaque répondant le jour de l'interview pour l'ENM de 2011, c'est-à-dire le 10 mai 2011. Par conséquent, pour les répondants de l'ECI de 2012, les utilisateurs doivent savoir que dans certains cas, il se peut que le répondant ait déménagé, que des changements soient survenus dans la composition de son ménage ou qu'il ait changé d'emploi entre la date de l'interview de l'ENM de 2011 et celle de l'interview de l'ECI de 2012. Autrement dit, une partie des renseignements provenant de l'ENM ne correspond peut-être pas à la situation du répondant au moment où l'interview de l'ECI a eu lieu.

La liste complète des variables dérivées de l'ENM accompagnées de leurs spécifications se trouve dans les dictionnaires de données de l'ECI de 2012.

2.7 Régions géographiques

L'Enquête canadienne sur l'incapacité (ECI) de 2012 a été conçue pour produire des données fiables pour chaque province et territoire. On peut également avoir accès à d'autres variables géographiques dans la base de données de l'ECI de 2012, dérivées des régions géographiques établies pour l'Enquête nationale auprès des ménages, comme les régions métropolitaines de recensement (RMR). De plus, les régions géographiques comprendront les régions sociosanitaires au Canada, qui représentent les secteurs administratifs ou régions administratives au sens où les autorités sanitaires l'entendent. Cependant, les utilisateurs doivent prendre note que ce ne sont pas toutes les données de l'ECI qui peuvent être croisées ou analysées à ces niveaux géographiques plus détaillés. Il sera possible de produire des tableaux de données, mais il y aura lieu d'examiner la fiabilité des estimations des données à ces niveaux géographiques en fonction de chaque cas.

Le dictionnaire de l'ENM présente les régions géographiques liées à l'ECI.

On trouvera de plus amples renseignements au sujet les régions sociosanitaires sur le site Web de Statistique Canada.

3. Plan d'échantillonnage de l'ECI

3.1 Population cible et couverture

La **population couverte** par l'ECI se compose de toutes les personnes âgées de 15 ans ou plus (en date du recensement/ENM, soit le 10 mai 2011), et qui présentent des limitations d'activités ou une restriction de participation liées à une condition physique ou mentale ou à un problème de santé, et qui vivaient au Canada au moment du recensement/ENM. Cette population inclut les personnes des ménages privés des 10 provinces et des 3 territoires. Cependant, pour des raisons d'ordre opérationnel, la population des réserves indiennes est exclue.

L'ECI tire son échantillon à partir des réponses fournies aux questions de l'Enquête nationale auprès des ménages (ENM) de 2011. Ainsi, toutes les personnes qui ont répondu « oui » à au moins une des questions filtres sur les limitations d'activité à l'ENM (voir Figure 1) ont été incluses dans la base de sondage de l'ECI. Ces personnes sont en effet très susceptibles d'avoir une incapacité, et font partie de la population qu'on appelle la **population des OUI**. Parmi ces personnes, celles dont les activités quotidiennes sont limitées en raison d'une condition ou d'un problème de santé à long terme font partie de la **population cible** des personnes avec incapacité.

Bien que l'ECI ne couvre pas la population ayant répondu négativement aux questions filtres de l'ENM (dite la **population des NON**), un échantillon de ces personnes est inclus dans le fichier analytique de l'ECI (voir section 3.8). Ces personnes sont toutes considérées comme étant des personnes sans incapacité. On retrouvera également un certain nombre de personnes sans incapacité dans la population des OUI. Elles sont représentées par les personnes qui n'ont pas déclaré de limitation d'activités au module des QII de l'ECI. Comme on le verra plus loin, l'échantillon de personnes sans incapacité sert à deux fins : le calcul de taux d'incapacité et la comparaison des caractéristiques de l'ENM des personnes avec et sans incapacité.

Figure 3.1
2011 NHS filter questions on activity limitations

ACTIVITÉS DE LA VIE QUOTIDIENNE

Enquête nationale auprès des ménages – N1 – question 7
Cette personne a-t-elle une quelconque **difficulté** à entendre, à voir, à communiqué, à marcher, à monter un escalier, à se pencher, à apprendre ou à faire d'autres activités semblables?

Oui, parfois
Oui, souvent
Non

Enquête nationale auprès des ménages – N1 – question 8
Est-ce qu'un état physique **ou** un état mental **ou** un problème de santé **réduit la quantité ou le genre d'activités** que cette personne peut faire :

(a) à la maison?

Oui, parfois
Oui, souvent
Non

(b) au travail ou à l'école?

Oui, parfois
Oui, souvent
Non
Ne s'applique pas

(c) dans d'autres activités, par exemple, dans les déplacements ou les loisirs?

Oui, parfois
Oui, souvent
Non

Source: Statistics Canada, 2011 National Household Survey.

3.2 Période de référence

L'Enquête canadienne sur l'incapacité (ECI) représente la population âgée de 15 ans ou plus en date du recensement/de l'ENM, soit le **10 mai 2011**. Cependant, étant donné le délai entre le Recensement/l'ENM et la collecte de l'ECI, toutes les informations recueillies lors de l'ECI représentent leurs caractéristiques telles que mesurées à **l'automne 2012**. Pour bien comprendre comment ces deux périodes de référence affectent l'utilisation et l'interprétation des données de l'ECI, voir la section 6.2.

3.3 Base de sondage tirée de l'ENM

La base de sondage a été construite à partir de la base de données de réponse du recensement et de l'ENM (communément appelée la BDR). Cette base contenait l'ensemble des réponses obtenues des divers modes de déclaration (Internet, questionnaires papiers, entrevues personnelles, etc.).

Plusieurs étapes de traitement ont été nécessaires sur la base de sondage afin que toutes les informations soient le plus complètes possible. Ainsi on a :

- imputé l'âge lorsque celui-ci était manquant sur le questionnaire de l'ENM, en se fiant sur les réponses au questionnaire de l'ENM (par exemple, données sur la scolarité, le travail, le revenu, etc.);
- utilisé différentes sources de données pour détecter des erreurs dans les numéros de téléphone ou pour trouver des numéros de téléphone ou adresses manquants;
- dérivé des poids préliminaires pour l'ENM afin de pouvoir estimer les tailles de population dans chaque strate lors de la répartition de l'échantillon.

3.4 Plan d'échantillonnage

Le plan de sondage de l'ECI peut être vu comme un plan à trois phases, où les deux premières phases correspondent à la sélection de l'échantillon de l'ENM lui-même, et la troisième phase correspond à la sélection du sous-échantillon de l'ECI.

Phase 1

La première phase correspondait à la sélection de l'échantillon de l'ENM lui-même. Le questionnaire de l'ENM comprend deux versions principales, le formulaire N1 et le formulaire N2.

Le formulaire N1 est complété par auto-dénombrement et est administré à environ un logement sur trois dans la plupart des régions du Canada (régions N1). En plus des questions démographiques de base du recensement (nom, sexe, date de naissance, état matrimonial légal, statut d'union libre, relation à la Personne 1, diverses questions de langue et la question portant sur le consentement à ce que les données soient diffusées dans 92 ans), le formulaire N1 de l'ENM comprend également des questions sur l'activité sur le marché du travail, le revenu, l'éducation, la citoyenneté, le logement, l'origine ethnique et ainsi de suite.

Le formulaire N2 est identique au formulaire N1, mis à part certaines questions exclues[2] et des exemples adaptés, et est administré par entrevue personnelle à tous les logements dans les régions éloignées et dans les réserves indiennes (régions N2).

Dans la terminologie d'échantillonage, la première phase du plan de sondage de l'ENM est un échantillon systématique stratifié de logements privés occupés avec une fraction de sondage constante de 1/3 dans les régions N1 et une énumération complète dans les régions N2.

2 Les questions sur la citoyenneté (question 10), le statut d'immigrant reçu (question 11) et l'année d'immigration (question 12) ne sont pas demandées aux personnes vivant dans les réserves indiennes et « settlements » énumérées à l'aide du formulaire 2.

Phase 2

La non-réponse de l'ENM n'était pas sans conséquence, alors une nouvelle stratégie a été mise en place pour diminuer l'effet potentiel d'un biais dû à la non-réponse. Ainsi un sous-échantillon de ménages non répondants a été sélectionné pour le suivi de la non-réponse (SNR). Par conséquent, ce sous-échantillon possède de plus grands poids de sondage pour réfléter les probabilités de sélection à chaque phase. Le sous-échantillon a été sélectionné de sorte à assurer une bonne représentation des unités de collecte (UC) avec plus forte concentration de certains groupes à risque de ne pas répondre[3]. Il est important de noter que le suivi de la non-réponse n'a pas été fait dans les régions N2. Ce sous-échantillon du SNR représente environ un ménage non répondant sur trois. Le plan de sondage de cette deuxième phase est un plan aléatoire assisté d'un modèle.

Phase 3

À la troisième phase, on a sélectionné l'échantillon de l'ECI parmi les personnes ayant répondu à l'ENM (incluant le sous-échantillon du SNR) et qui ont déclaré des limitations d'activités à l'ENM. Sont exclues de la sélection les personnes vivant sur des réserves indiennes et les personnes de moins de 15 ans au 10 mai 2011.

L'échantillon de l'ECI a été sélectionné de manière à disposer d'échantillons suffisamment importants dans chaque domaine d'estimation, tel qu'expliqué à la section 3.5.

3.5 Domaines d'estimation et stratification

Les domaines d'estimation désignent les groupes d'unités pour lesquels des estimations doivent être produites avec un degré de précision « acceptable ». Ces domaines d'estimation pour l'ECI correspondent au croisement de chaque province par les groupes d'âge ci-dessous :

- de 15 à 24 ans
- de 25 à 44 ans
- de 45 à 64 ans
- de 65 à 74 ans
- 75 ans ou plus

Pour chacun des trois territoires, le domaine d'estimation comprend un seul groupe d'âge (15 ans et plus). À l'Île-du-Prince-Édouard, les deux premiers groupes d'âge ont dû être combinés puisque les tailles de population dans ces groupes d'âge étaient très petites.

Chaque domaine d'estimation a ensuite été subdivisé en strates, ce qui permet de tenir compte de la sévérité[4] de la limitation d'activités ainsi que du plan de sondage de l'ENM. Pour neutraliser la sévérité, les personnes qui ont répondu « Oui, souvent » au moins une fois dans les questions filtres de l'ENM ont été placées dans la strate « Souvent » ; par ailleurs, les personnes qui ont répondu « Oui, parfois » au moins une fois, mais jamais « Oui, souvent » ont été rangées dans la strate « Parfois ». Pour tenir compte du plan d'échantillonnage de l'ENM, nous avons pris deux caractéristiques en considération : le fait de figurer au nombre du premier groupe de répondants de l'ENM (strates des premiers répondants) par opposition aux répondants qui ont été l'objet d'un suivi de la non-réponse (les strates SNR) et vivant dans une région éloignée (les strates N2) ou non (les strates N1). Notez qu'aucun suivi de non-réponse n'était effectué dans les régions éloignées. En tenant compte de ses caractéristiques, il a été possible de mieux regrouper les gens avec des poids provisoires semblables.

3 De nombreuses simulations ont été effectuées en préparation pour la collecte de l'ENM afin de tenter d'identifier les groupes les plus à risque de ne pas répondre à l'enquête. Les UC comprenant de fortes concentrations de ces groupes cibles ont donc été sur-échantillonnées pour le suivi de la non-réponse.

4 Le concept de "sévérité" dont il est question ici n'est pas le même que celui utilisé pour le score de sévérité décrit à la section 2.3. Ici, on tient compte uniquement de la fréquence des limitations rapportées aux questions filtres de l'ENM.

Ainsi, chaque domaine d'estimation était divisé en 6 strates définies comme suit :

- région N1 – répondant initial - « souvent »
- région N1 – répondant initial - « parfois »
- région N1 – répondant au SNR - « souvent »
- région N1 – répondant au SNR - « parfois »
- région N2 – répondant initial - « souvent »
- région N2 – répondant initial - « parfois »

Notez que ces 6 strates n'étaient pas toujours présentes dans chaque domaine d'estimation. Toutes les personnes rencontrant les conditions pour faire partie de la base de sondage de l'ECI étaient ainsi classifiées dans les domaines d'estimation et strates avant de procéder à la sélection de l'échantillon.

3.6 Répartition de l'échantillon

Les unités d'échantillon ont été réparties de sorte que, pour chaque domaine d'estimation, on puisse estimer une proportion minimale avec un coefficient de variation (CV) maximal de 16,5 %. À Statistique Canada, on utilise souvent la valeur de 16,5 % comme limite supérieure d'un CV permettant de qualifier l'estimation correspondante d'acceptable. Le tableau suivant donne les proportions minimales à estimer dans chaque domaine d'estimation. Pour ces calculs, on a supposé un effet de plan de 1,2. C'est-à-dire qu'on a supposé que dans les domaines d'estimation, la variance qui serait obtenue avec le plan de sondage de l'ECI serait 20 % plus grande que celle qu'on obtiendrait si on tirait un échantillon aléatoire simple de même taille dans chaque domaine.

Tableau 3.1
Proportions minimales à estimer avec un coefficient de variation (CV) maximal de 16,5% par domaine d'estimation

Province/Territoire	Groupes d'âge				
	15 à 24 ans	25 à 44 ans	45 à 64 ans	65 à 74 ans	75 ans et plus
			%		
Terre-Neuve-et-Labrador	9,0	8,0	8,0	11,0	11,0
Île-du-Prince-Édouard	9,0		8,0	11,0	11,0
Nouvelle-Écosse	9,0	8,0	8,0	11,0	11,0
Nouveau-Brunswick	9,0	8,0	8,0	11,0	11,0
Québec	9,0	8,0	8,0	11,0	11,0
Ontario	9,0	8,0	8,0	11,0	11,0
Manitoba	9,0	8,0	8,0	11,0	11,0
Saskatchewan	9,0	8,0	8,0	11,0	11,0
Alberta	9,0	8,0	8,0	11,0	11,0
Colombie Britannique	9,0	8,0	8,0	11,0	11,0
Yukon			8,0		
Territoires du Nord-Ouest			8,0		
Nunavut			8,0		

Source : Statistique Canada, Enquête canadienne sur l'incapacité, 2012.

Une méthode de répartition optimale entre les strates d'un domaine particulier a été utilisée en tenant compte de la non-réponse attendue et du taux de faux positifs attendus (les personnes ayant déclaré des limitations d'activités à l'ENM mais qui n'ont pas d'incapacité selon les QII) dans chacune. Cette répartition dépendait entre autres des poids de l'ENM corrigés pour la non-réponse. Il est à noter qu'au moment de faire la répartition, ces poids n'avaient pas encore été calculés. Des poids préliminaires ont donc été

calculés aux fins de cette répartition seulement. Toutefois les poids finaux de l'ENM ont été utilisés dans le processus de pondération de l'ECI. Pour une référence au sujet de la pondération de l'ENM, voir Morel et Nambeu (2013)[5].

Suite à la répartition de l'échantillon, il arrivait qu'on ait besoin de plus d'unités dans une strate que le nombre disponible sur la base de sondage. Dans ce cas, on devait sélectionner toutes les unités de cette strate (en faire une strate à tirage complet) et sélectionner le nombre d'unités en déficit dans les autres strates du même domaine d'estimation proportionnellement à leur répartition dans celles-ci.

3.7 Tailles d'échantillons

Le tableau 3.2 ci-dessous donne les tailles d'échantillon finales de l'échantillon de l'ECI (échantillon des OUI) par province et territoire.

Tableau 3.2
Tailles d'échantillon par province et territoire

Province/Territoire	Échantillon envoyé à la collecte
Terre-Neuve-et-Labrador	3 792
Île-du-Prince-Édouard	2 809
Nouvelle-Écosse	3 769
Nouveau-Brunswick	4 146
Québec	4 254
Ontario	4 563
Manitoba	5 081
Saskatchewan	4 372
Alberta	4 740
Colombie Britannique	5 197
Yukon	913
Territoires du Nord-Ouest	836
Nunavut	971
TOTAL	**45 443**

Source : Statistique Canada, Enquête canadienne sur l'incapacité, 2012.

3.8 Échantillon des NON

Tel que mentionné précédemment, le fichier analytique de l'ECI sert entre autres à estimer des taux d'incapacité à divers niveaux géographiques, mais aussi à comparer les caractéristiques des personnes avec incapacité avec celles des personnes sans incapacité. Pour ce faire, le fichier analytique doit également inclure un échantillon représentatif des personnes sans incapacité, ce qui n'est pas le cas si on considère seulement l'échantillon des OUI. On a donc aussi sélectionné un échantillon à partir de la population des NON, soit les personnes n'ayant pas déclaré de limitations d'activités aux questions filtres de l'ENM. L'hypothèse sous-jacente ici est que les personnes appartenant à la population des NON à l'ENM sont moins susceptibles d'avoir une incapacité, ou sinon que cette incapacité est très légère. La décision de ne pas couvrir la population des NON à l'ECI est aussi reliée au coût que cela comporte, puisqu'il faut un très grand échantillon de cette population pour arriver à trouver des personnes avec incapacité, ce qui est très coûteux lors de la collecte.

L'échantillon des NON n'a pas à être envoyé sur le terrain puisque chaque personne est considérée comme n'ayant pas d'incapacité. Dans le fichier analytique, on retrouve un grand nombre de caractéristiques provenant de l'ENM pour cet échantillon, de même que pour l'échantillon des OUI, ce qui permet de faire la comparaison des personnes avec et sans incapacité. Grâce à cet échantillon de NON, il est aussi possible pour les analystes d'obtenir les dénominateurs requis au calcul des taux d'incapacité pour plusieurs sous-groupes

5 Morel, J. et Nambeu, C. (2013). «National Household Survey: Weighting and estimation update». Présenté à la conférence ACS en mai 2013.

de la population canadienne. Pour éviter que le fichier analytique soit trop gros, on a tiré un échantillon de la population des NON plutôt que de prendre toute la population. La taille de cet échantillon a été fixée de sorte à obtenir la même précision que celle demandée pour l'échantillon des OUI (voir tableau 3.1), avec un CV maximal de 7%[6]. La sélection de l'échantillon a été faite en stratifiant la population des NON selon la province, le sexe, et des groupes d'âges de 5 ans, le dernier groupe étant les 75 ans et plus (donc plus détaillé que pour l'échantillon des OUI). Le plan de sondage de l'ENM a également été pris en compte pour l'échantillon des NON puisque les strates ont ensuite été sous-divisées selon la région N1 ou N2, et selon le fait d'être un répondant initial ou un répondant au suivi de la non réponse. Le tableau suivant donne les tailles d'échantillon par province.

Tableau 3.3
Tailles d'échantillon des NON, par province et territoire

Province/Territoire	Échantillon des NON (non envoyés à la collecte)
Terre-Neuve-et-Labrador	11 813
Île-du-Prince-Édouard	6 643
Nouvelle-Écosse	12 062
Nouveau-Brunswick	11 997
Québec	12 455
Ontario	12 486
Manitoba	12 152
Saskatchewan	12 103
Alberta	12 338
Colombie Britannique	12 426
Yukon	2 499
Territoires du Nord-Ouest	2 578
Nunavut	2 446
TOTAL	**123 998**

Source : Statistique Canada, Enquête canadienne sur l'incapacité, 2012.

Puisque toutes les unités de la population des NON sont considérées comme n'ayant pas d'incapacité, le score de sévérité de chaque unité dans le fichier a été fixé à 0, et la classe de sévérité a été fixée à 0.

6 For Prince Edward Island, we set the CV at 10% for the 65-to-74 age group and 15% for the 75-and-over age group to avoid selecting all available units in the sampling frame.

4. Collecte des données

4.1 Période de référence

L'Enquête canadienne sur l'incapacité (ECI) s'est déroulée du 24 septembre 2012 au 13 janvier 2013. À titre d'enquête postcensitaire ou post-ENM, elle suivait l'Enquête nationale auprès des ménages de 2011, qui avait eu lieu le 10 mai 2011. Il s'était écoulé un délai de 16 à 20 mois entre les deux enquêtes.

4.2 Communications

Dans les mois précédant la collecte des données pour l'Enquête canadienne sur l'incapacité (ECI), on a organisé des activités promotionnelles afin de sensibiliser le public à l'enquête et à encourager la participation.

Statistique Canada a ajouté sur son site Web une page où figurait une liste de questions et de réponses à l'intention des répondants, une description de l'enquête, des renseignements généraux sur l'enquête et sa méthodologie ainsi que le lien donnant accès au questionnaire. Statistique Canada a conçu une icône spéciale pour son site Web, qui' a été distribuée aux membres du Groupe consultatif technique sur les personnes ayant une incapacité d'EDSC. Ceux-ci ont contribué à faire la promotion de l'enquête en affichant l'icône de l'enquête sur le site Web de leur organisation, donnant ainsi accès à la page Web de Statistique Canada sur l'ECI.

Avant la collecte des données, Statistique Canada a envoyé aux répondants une lettre de présentation et un dépliant afin de les mettre au courant de l'enquête à venir et de leur faire valoir l'importance de leur participation à celle-ci. La lettre a été rédigée dans les deux langues officielles. Pour la collecte au Nunavut, la lettre a été également distribuée en inuktitut (dialecte parlé sur l'île de Baffin). De plus, Statistique Canada a joint un encart en braille aux lettres de présentation, contenant le numéro de téléphone d'une personne-ressource pour que les personnes ayant une déficience visuelle puissent communiquer avec Statistique Canada pour obtenir de plus amples renseignements.

4.3 Mode de collecte

Les questions de l'Enquête canadienne sur l'incapacité (ECI) de 2012 ont été administrées sous la forme d'une interview téléphonique assistée par ordinateur (ITAO). On a pris des mesures d'adaptation afin d'optimiser la participation directe des personnes ayant une incapacité en offrant à tous les répondants une adresse de courriel, un numéro de téléphone ATS (appareil de télécommunication pour malentendants) et un numéro de téléphone d'une personne-ressource sur l'encart en braille accompagnant la lettre de présentation. Dans les cas où les répondants demandaient d'autres arrangements, on a procédé à des interviews sur place avec la méthode des interviews utilisant papier et crayon (IPC).

On s'est également servi de la méthode des IPC pour mener des interviews en personne auprès de certains répondants des Territoires du Nord-Ouest qui n'avaient pas de téléphone et qu'il aurait été impossible de contacter autrement. Avant l'enquête, on savait déjà qu'un fort pourcentage des logements des territoires de l'Ouest n'avaient pas de téléphone, ce qui impliquait que la couverture se limiterait généralement aux répondants de l'ENM qui avaient un téléphone. Avant de commencer la collecte des données, Statistique Canada a réussi à conclure un partenariat avec le bureau de la statistique des Territoires du Nord-Ouest afin de créer une équipe d'intervieweurs chargés d'aider à mener les interviews sur place. On a donc constitué un échantillon de logements qui n'avaient pas de téléphone et utilisé la méthode des IPC pour recueillir les données en faisant du porte-à-porte.

Partout au Canada, les interviews se sont déroulées dans la langue officielle choisie par le répondant (le français ou l'anglais). Le questionnaire a également été traduit en inuktitut, et les interviews dans cette langue ont été mises à la disposition des répondants vivant au Nunavut.

Le temps nécessaire pour répondre au questionnaire variait d'une personne à l'autre, mais, en moyenne, les répondants ont pris une quarantaine de minutes pour le faire.

4.4 Supervision et contrôle de la qualité

Tous les intervieweurs de Statistique Canada ont travaillé sous la supervision d'intervieweurs principaux qui avaient pour responsabilité de s'assurer que les intervieweurs connaissaient bien les concepts et les procédures associés aux enquêtes auxquelles ils étaient affectés. Les intervieweurs principaux étaient également chargés d'observer périodiquement les intervieweurs afin de s'assurer qu'ils suivaient les procédures courantes.

Les intervieweurs ont reçu une formation sur le contenu de l'enquête et l'application d'interview assistée par ordinateur. En plus des exposés en classe et d'une période d'autoformation, les intervieweurs ont mené une série d'interviews simulées afin de se familiariser avec l'enquête ainsi que ses concepts et ses définitions.

4.5 Interviews par procuration

L'incapacité étant difficile à mesurer et très subjective, on demandait aux intervieweurs de faire tous les efforts possibles pour mener l'interview auprès de la personne choisie. Cependant, dans les circonstances suivantes, une interview par procuration pouvait être acceptée :

- La personne choisie était absente pendant l'enquête.
- La personne choisie ne parlait ni anglais ni français.
- La personne choisie ne pouvait pas participer en raison de problèmes de santé mentale ou physique.
- Un parent insistait pour répondre pour son enfant âgé de 15 à 17 ans.

Le candidat substitut devait satisfaire aux conditions suivantes afin d'être accepté :

- Être un adulte et parler soit le français ou l'anglais;
- Pouvoir être joint pendant la collecte des données de l'enquête;
- Être la personne la mieux renseignée, ou parmi les mieux renseignées sur les difficultés liées aux limitations d'activités et aux restrictions à la participation de la personne choisie.

Au total, 5 164 entrevues ont été faites par procuration et de celles-ci, 93 % ont été complétées (i.e. avec un statut final de « répondant »).

Le tableau suivant donne la distribution des entrevues par procuration selon l'âge du répondant et la raison de l'entrevue par procuration (excluant 320 personnes pour qui on ne connaît pas la raison).

Tableau 4.1
Distribution du nombre d'entrevues par procuration selon l'âge du répondant et la raison de l'entrevue par procuration

Groupes d'âge	Raison de l'entrevue par procuration				
	Santé	Absent	Langue	Parent* insiste pour répondre	TOTAL
15 à 24 ans	835	431	30	328	1 624
25 à 44 ans	360	150	100	8	618
45 à 64 ans	235	132	167	7	541
65 à 74 ans	335	62	178	5	580
75 ans et plus	863	47	216	7	1 133
TOTAL	**2 628**	**822**	**691**	**355**	**4 496**

*Ici, le mot "parent" est utilisé au sens large. Il s'agit souvent du parent ou du tuteur lorsqu'il est question d'un jeune répondant mais il pourrait s'agir du conjoint ou d'un autre membre de la famille pour les répondants plus âgés.
Source : Statistique Canada, Enquête canadienne sur l'incapacité, 2012.

4.6 Problèmes particuliers

Les intervieweurs ont reçu comme consigne de faire toute tentative raisonnable pour obtenir une interview complète avec le membre du ménage qui avait été choisi. On a communiqué de nouveau, et ce, jusqu'à deux reprises, avec les personnes qui avaient d'abord refusé de participer, afin de leur expliquer l'importance de l'enquête et de les inciter à y participer. Dans les cas où l'intervieweur avait téléphoné à un mauvais moment, il prenait rendez-vous en vue de rappeler à un moment plus opportun. Dans les cas où il n'y avait personne à la maison, de nombreux appels ont été faits.

Des problèmes particuliers ont surgi par rapport à la collecte de données pour l'Enquête canadienne sur l'incapacité (ECI), problèmes qui ont été abordés avec plus de coordination sur le terrain et par des corrections aux méthodes de l'enquête. Par exemple, en raison du long délai qui s'était écoulé entre l'ENM de 2011 et l'ECI de 2012, il est arrivé que les coordonnées de certains répondants soient inexactes. Pour cette raison, il a été difficile de livrer les lettres de présentation à certains répondants qui ne vivaient plus à l'adresse qu'ils avaient indiquée à l'époque où l'ENM 'a eu lieu. Pour ce qui est des numéros de téléphone qui n'étaient plus en fonction, les intervieweurs avaient reçu des directives spéciales sur la façon de retracer les répondants afin de procéder à l'interview.

Comme ça arrive parfois, la collecte des données de l'ECI de 2012 a lieu en même temps que celle de plusieurs autres enquêtes menées par Statistique Canada, ce qui risquait de représenter un fardeau pour les répondants. On a planifié soigneusement la conception de l'enquête et on y a apporté des ajustements afin de réduire le fardeau pour les ménages et les répondants au questionnaire.

La description complète des ajustements apportés aux poids de l'enquête à la suite de problèmes particuliers qui ont surgi pendant la collecte des données est présentée à la section 6.1.

4.7 Taux de réponse

La collecte de l'ECI s'est terminée avec un taux de réponse de 74.6 %. Ce taux de réponse est calculé comme étant le nombre de répondants complets (avec ou sans incapacité), divisé par le nombre de cas envoyés à la collecte duquel on a soustrait les hors-champs. Les cas « hors du champ de l'enquête » comprennent les gens qui sont décédés, qui ont émigré, qui ont été admis dans un établissement institutionnel, qui ont déménagé dans une réserve indienne, qui font partie des Forces canadiennes, qui sont des visiteurs au Canada (classés par erreur pendant l'ENM) ou dont l'âge est invalide. Ce taux reflète le pourcentage de cas qui ont complété l'entrevue par rapport au nombre qui devaient la compléter (d'où l'exclusion des hors champs au dénominateur).

Taux de réponse = Cas complétés / (Cas envoyés à la collecte – Cas hors champs)

Les prochains tableaux donnent des taux de réponse par province et territoire et par groupe d'âge.

Tableau 4.2
Taux de réponse par province et territoire

Province/Territoire	Envoyés à la collecte	Hors champs	Complétés	Taux de réponse
	nombre			%
Terre-Neuve-et-Labrador	3 792	156	2 733	75,2
Île-du-Prince-Édouard	2 809	118	2 126	79,0
Nouvelle-Écosse	3 769	141	2 819	77,7
Nouveau-Brunswick	4 146	109	2 761	68,4
Québec	4 254	134	3 182	77,2
Ontario	4 563	117	3 243	72,9
Manitoba	5 081	158	3 645	74,0
Saskatchewan	4 372	154	3 356	79,6
Alberta	4 740	119	3 522	76,2
Colombie Britannique	5 197	122	3 635	71,6
Yukon	913	20	655	73,3
Territoires du Nord-Ouest	836	35	545	68,0
Nunavut	971	37	600	64,2
Total	45 443	1 420	32 822	74,6

Source : Statistique Canada, Enquête canadienne sur l'incapacité, 2012.

Tableau 4.3
Taux de réponse par groupe d'âge

Groupes d'âge	Envoyés à la collecte	Hors champs	Complétés	Taux de réponse
	nombre			%
15 à 24 ans	9 944	121	6 633	67,5
25 à 44 ans	11 672	119	7 831	67,8
45 à 64 ans	10 487	180	8 182	79,4
65 à 74 ans	6 481	244	5 282	84,7
75 ans et plus	6 859	756	4 894	80,2

Source : Statistique Canada, Enquête canadienne sur l'incapacité, 2012.

5. Traitement des données

5.1 Saisie des données

Les réponses au questionnaire de l'enquête ont été saisies directement par l'intervieweur dans le questionnaire automatisé. Il s'agissait d'une application mise au point au moyen du logiciel Blaise dans un système d'interviews téléphoniques assistées par ordinateur (ITAO). L'emploi d'un questionnaire automatisé réduit les délais et les coûts de traitement associés à la saisie des données, aux erreurs de transcription et à la transmission des données. Les données recueillies à l'aide du questionnaire papier de l'ECI ont également été entrées dans le système d'ITAO après leur retour du terrain et elles ont donc été saisies par voie électronique pour être traitées de façon plus poussée.

Une partie du contrôle se fait au moment de l'interview. Plus particulièrement, lorsqu'une réponse donnée ne semble pas correspondre aux réponses antérieures ou qu'elle tombe à l'extérieur des valeurs prévues, le système demande à l'intervieweur, en affichant des messages à l'écran, de confirmer les réponses auprès du répondant, et, au besoin, de modifier l'information. On soumet ensuite les données des réponses à d'autres processus de vérification et d'imputation après réception au bureau central.

5.2 Étapes du traitement

Le traitement des données comporte une série d'étapes pour la conversion des réponses au questionnaire, de leur format brut initial en une base de données conviviale de grande qualité, contenant un vaste ensemble de variables à analyser. L'équipe du traitement exécute une série d'opérations avec les données afin d'enlever des fichiers les erreurs commises par inadvertance, de modifier les données aux fins de cohérence, de coder les questions à développement, de créer les variables utiles aux fins de l'analyse des données et enfin de systématiser et de documenter les variables pour en faciliter l'analyse.

Pour l'ECI de 2012, nous nous sommes servis d'une nouvelle série d'outils de traitement des enquêtes sociales appelée « Environnement pour le traitement des enquêtes sociales » (ETES). L'ETES comprend des programmes du logiciel SAS, des applications sur mesure et des procédures manuelles permettant d'exécuter les étapes de traitement systématique suivantes :

- Réception des données brutes
- Nettoyage
- Recodage
- Enchaînement des questions
- Codage
- Contrôle et imputation
- Variables dérivées
- Création du fichier de traitement final
- Création des fichiers de diffusion

5.3 Nettoyage des enregistrements : enregistrements dans le champ de l'enquête et enregistrements complets

Nous avons mis en place un certain nombre de procédures de nettoyage préliminaire pour l'ECI de 2012 au niveau des enregistrements individuels, que l'on utilise après avoir reçu les données brutes recueillies à l'aide du questionnaire électronique. L'une de ces mesures rigoureuses employées systématiquement pour protéger la confidentialité des répondants consistait en la suppression dans les fichiers tous les renseignements personnels, comme le nom et l'adresse. C'est à cette étape que les enregistrements en double étaient corrigés. De plus, au nombre des procédures de nettoyage figurait l'examen des enregistrements de tous les répondants, ce qui permettait de vérifier si chaque répondant était « dans le champ de l'enquête » et s'il avait répondu à un nombre suffisant de questions pour avoir un questionnaire « complet ». Voici les critères que les répondants devaient remplir :

1. Pour être « dans le champ de l'enquête » pour l'ECI de 2012, les répondants devaient être âgés d'au moins 15 ans le jour de l'Enquête nationale auprès des ménages (ENM), le 10 mai 2011. Les répondants qui sont « dans le champ de l'enquête » se rangent dans deux groupes : 1) les répondants qui ont été sélectionnés après avoir répondu aux Questions d'identification des incapacités (QII) et qui faisaient donc partie de la population ayant une incapacité; 2) les répondants qui n'ont pas été retenus après avoir répondu aux QII et qui étaient donc considérés comme des personnes sans incapacité. Les deux groupes se retrouvent dans la base de données de l'enquête.
2. Pour avoir un questionnaire « complet », les répondants qui remplissaient les critères de la population ayant une incapacité devaient avoir répondu au nombre minimum de questions essentielles dans l'enquête, nécessaires à la production des tableaux de données sur les personnes ayant une incapacité selon les exigences stipulées par la Loi sur l'équité en matière d'emploi de 1995. Ces questions figurent à l'annexe C.
3. Pour avoir un questionnaire « complet », les répondants qui avaient été classés dans la population sans incapacité ne devaient avoir répondu qu'aux questions des QII. Le reste du questionnaire ne s'appliquait pas à eux.

Les répondants qui ne remplissaient pas ces critères ont été retirés de la base de données et ils ont été classés dans la catégorie des non-répondants.

5.4 Recodage des variables et questions à réponses multiples

Cette étape du traitement consistait en des modifications de variables individuelles. Certaines variables pouvaient être éliminées, recodées, redimensionnées ou laissées telles quelles. On a apporté des changements au formatage dans le but de faciliter le traitement ainsi que l'analyse des données par les utilisateurs finaux. On a, entre autres choses, laissé tomber la lettre Q qui apparaissait au début du numéro de chaque question (voir ci-dessous).

Parmi les autres changements apportés aux variables, on note la transformation des questions à réponses multiples (questions pour lesquelles toutes les réponses appropriées doivent être choisies) en séries correspondantes de variables à réponse unique, qui sont plus faciles à utiliser. À chaque catégorie de réponses associée à la question initiale, on a créé une variable avec « oui/non » comme valeur de réponse. Reportez-vous à l'exemple ci-après. Cette procédure s'appelle « désenchaîner » les variables et elle facilite la mise en tableaux statistiques des variables par les utilisateurs finaux.

Question initiale à réponses multiples :

DRV_Q06 Quelles sont les raisons pour lesquelles vous avez de la difficulté à utiliser les modes de transport en commun ou le service de transport adapté?

INTERVIEWEUR : Exemples de services de transport adapté : Handi-Transit, Wheel-Trans, Para-Transpo, etc.

Choisissez toutes les réponses appropriées. Lisez les catégories au répondant.

01 Service non disponible lorsque vous en avez besoin
02 Les modalités de réservation ne permettent pas les arrangements de dernière minute
03 Difficulté à vous rendre aux arrêts d'autobus ou à repérer les points d'arrêts d'autobus
04 Difficulté à monter à bord du véhicule ou à en descendre
05 Difficulté à voir les affiches, les avis, les arrêts ou à entendre les annonces
06 Réseau achalandé/transports en commun bondés
07 Difficulté à demander les services
08 Difficulté à interpréter les horaires
09 Correspondance difficile à effectuer ou compliquée
10 Sortir aggrave votre état
11 Trop dispendieux
12 Autre raison
NSP, RF

Variables définitives sous la forme de questions à réponse unique « oui/non » :

DRV_06A Quelles sont les raisons pour lesquelles vous avez de la difficulté à utiliser les modes de transport en commun ou le service de transport adapté?

- Service non disponible lorsque vous en avez besoin

1 Oui
2 Non
NSP, RF

DRV_06B Quelles sont les raisons pour lesquelles vous avez de la difficulté à utiliser les modes de transport en commun ou le service de transport adapté?

- Les modalités de réservation ne permettent pas les arrangements de dernière minute

1 Oui
2 Non
NSP, RF

DRV_06C Quelles sont les raisons pour lesquelles vous avez de la difficulté à utiliser les modes de transport en commun ou le service de transport adapté?

- Difficulté à vous rendre aux arrêts d'autobus ou à repérer les points d'arrêts d'autobus

1 Oui
2 Non
NSP, RF

...autres questions dont la réponse est « Oui » ou « Non » pour chaque catégorie de réponses, selon les indications, de « Difficulté à monter à bord du véhicule ou à en descendre » à « Trop dispendieux »... y compris la dernière catégorie :

DRV_06L Quelles sont les raisons pour lesquelles vous avez de la difficulté à utiliser les modes de transport en commun ou le service de transport adapté?

- Autre raison

1 Oui
2 Non
NSP, RF

Dans certains cas, les questions à réponses multiples dans l'enquête ont des questions de suivi correspondantes, qui sont également des questions à réponses multiples. Voilà pourquoi, une seule question peut devenir un large éventail de variables dans la base de données finale.

Par exemple, **AAD_Q10** se lit comme suit : **Quels sont les aides ou les appareils fonctionnels dont vous avez besoin, mais que vous n'avez pas?** En tout, 17 catégories de réponse pour les appareils fonctionnels sont présentées (et une catégorie supplémentaire a été créée lors du du codage (voir la discussion ci-dessous), soit 18 catégories finales en tout). Par conséquent, **AAD_Q10** se transformera en 18 variables distinctes de type « oui/non » : **AAD_10A** à **AAD_10R**. La question est suivie de **AAD_Q11 – Pourquoi n'avez-vous pas (elle fait référence aux réponses de AAD_Q10)?** Pour cette question de suivi, 8 catégories de raisons sont présentées. Donc, pour **AAD_Q11**, une gamme de 144 (18 x 8) variables de type « oui/non » a été créée, dans laquelle les 18 appareils correspondent aux lettres A à R inclusivement, et les 8 raisons, aux lettres A à H inclusivement : **AAD_11AA, AAD_11AB, AAD_11AC... AAD_11AH** à **AAD_11RA, AAD_11RB, AAD_11RC... AAD_11RH**.

5.5 Enchaînement des questions : parcours des réponses, enchaînements valides et la non-réponse aux questions

La vérification de l'enchaînement des questions ou les instructions « passez à » représente une autre série de procédures pour le traitement des données de l'ECI de 2012. On a vérifié tous les parcours des réponses et toutes les instructions « passez à » intégrées au questionnaire afin de s'assurer que l'univers ou la population visée de chaque question soit saisie précisément au cours du traitement. On a accordé une attention toute particulière aux distinctions entre les enchaînements valides et la non-réponse, ce qui représente une distinction importante aux fins de l'analyse statistique. Ces concepts sont exposés ci-après, ce qui aidera les utilisateurs à mieux comprendre les univers de questions ainsi que les résultats statistiques pour les variables de l'ECI.

Réponse – Une réponse se rapportant directement au contenu de la question, qui peut être classée dans les catégories de réponses établies d'avance, y compris « Autre - Précisez ».

Enchaînement valide (saut valide) – Il s'agit d'une question que le répondant a sautée parce qu'elle ne s'appliquait pas à sa situation, compte tenu des réponses valides à des questions précédentes. Dans un tel cas, on considère que le répondant ne fait pas partie de la population visée (l'univers) de cette question. Comme nous l'avons fait remarquer ci-dessous, lorsque le répondant saute une question pour une raison indéterminée (c'est-à-dire que la réponse « Je ne sais pas » ou « Refus » à une question antérieure l'a mené à sauter la question), on lui attribue le code « Non déclaré » pour cette question.

Ne sais pas – Le répondant était incapable de répondre, et ce, pour différentes raisons (par exemple, il avait du mal à se souvenir ou il répondait à la place de quelqu'un d'autre).

Refus – Le répondant a préféré ne pas répondre, peut-être en raison de la nature délicate de la question.

Non déclaré – Cette catégorie indique que le répondant n'a pas répondu et qu'un parcours n'est pas déterminé pour ce dernier, comme lorsque le répondant a omis de répondre à la question filtre précédente ou que l'on a relevé une incohérence dans une série de réponses.

Supprimé – Pour les variables de l'ENM seulement, un type supplémentaire de non-réponse est la valeur supprimée en raison des procédures de contrôle de la cohérence ayant eu lieu au cours du traitement des données de l'ECI (voir la section 5.7).

On a affecté des codes spéciaux à chaque type de réponses, dans le but d'aider les utilisateurs à les reconnaître et à faciliter l'analyse des données. Par exemple, les codes « enchaînement valide » se terminent par un « 6 », précédé d'un ou de plusieurs « 9 » (par exemple, le code serait « 996 » pour les variables à trois chiffres). Pour toutes les réponses « Ne sais pas », le code se termine par un « 7 », précédé par un ou plusieurs « 9 » (par exemple, « 997 »). Pour les refus, le code se termine par un « 8 », précédé par un ou plusieurs « 9 » (par exemple, « 998 ») ; pour « Non déclaré », le code se termine par un 9, précédé par un « 9 » également (par exemple, « 999 »). Pour ce qui est des valeurs supprimées, on entre le code « 5 » comme dernier chiffre, et les chiffres précédents sont des « 9 » (par exemple « 995 »).

5.6 Codage

Réponse « Autre – Précisez »

Le traitement des données consiste également à coder les réponses « Autre – Précisez » (réponses écrites). Pour la plupart des questions de l'ECI, on avait prévu des catégories de réponses précodées, et les intervieweurs avaient reçu de la formation leur permettant de coder la réponse d'un répondant à la catégorie appropriée. Cependant, si la réponse d'un répondant était difficile à classer dans une catégorie existante, l'intervieweur pouvait également pour bien des questions entrer une réponse en toutes lettres dans la catégorie « Autre – Précisez ».

Toutes les questions dans les catégories « Autre – Précisez » ont été examinées au cours du traitement. Dans le cas de questions pour lesquelles les réponses « Autre – Précisez » représentent moins de 10 % du total des réponses à la question, le codage n'a pas été effectué, et les réponses sont demeurées dans la catégorie « Autre ». Dans le cas de12 éléments du questionnaire, on a codé les réponses « Autre – Précisez ». Selon les lignes directrices sur le codage, le code de beaucoup de réponses données en toutes lettres par les répondants à ces questions a été remplacé par le code de l'une des catégories de réponses existantes. Les réponses uniques et différentes sur le plan qualitatif par rapport aux catégories existantes sont demeurées dans la catégorie « Autre ». Lorsque les nombres le justifiaient, on a créé des catégories pour saisir les nouveaux thèmes qui ressortaient des données, que les catégories existantes ne reflétaient pas. L'annexe D présente les catégories supplémentaires ajoutées à l'ECI de 2012. Elles seront prises en considération au moment de revoir les catégories de réponse pour les cycles à venir de l'enquête.

Questions ouvertes et classifications types

Les intervieweurs ont enregistré quelques questions de l'ECI de 2012 dans un format complètement ouvert. Il s'agit des questions se rapportant à ce qui suit :

1. Le principal problème de santé du répondant, qui lui cause le plus de difficultés ou qui limite le plus ses activités;
2. Sa profession et le secteur d'activité dans lequel il travaille;
3. Le principal domaine d'études postsecondaires, le cas échéant.

On a codé ces réponses en se servant d'une combinaison de procédures de codage automatisées et interactives (manuelles). Les systèmes de classification type ont servi à coder ses réponses. Des précisions sur ces classifications figurent à l'annexe D.

Le codage pour toutes les classifications a fait appel à une équipe de codeurs chevronnés et de surveillants du contrôle de la qualité. Des spécialistes dans le domaine du traitement des données ont appliqué d'autres procédures de vérification.

5.7 Vérifications et imputations

L'étape du codage a été suivie d'analyses des données visant à relever les lacunes, les incohérences, les valeurs aberrantes et autres problèmes potentiels dans les données. Afin de régler les problèmes relevés dans les données, on a procédé à des vérifications et à des imputations personnalisées.

On a évalué les données de l'ECI afin de faire ressortir les incohérences internes entre les variables ainsi que les incohérences externes par rapport aux données de l'ENM couplées des répondants. Pour ce qui est des incohérences internes, on avait déjà soumis les données à d'importantes vérifications, intégrées dans le questionnaire électronique.

Mentionnons, à titre d'exemple de vérifications électroniques, plusieurs vérifications qui consistaient à s'assurer que les réponses se rapportant à l'âge dans le questionnaire ne se contredisaient pas. Par exemple, certaines vérifications consistaient à s'assurer que l'âge indiqué par le répondant au moment où le répondant s'est rendu compte pour la première fois d'une limitation dans ses activités quotidiennes ou lorsqu'il a obtenu son niveau de scolarité le plus élevé ne dépassait pas l'âge du répondant au moment de l'interview. Dans une autre vérification du questionnaire : il s'agissait de s'assurer que le nombre d'heures

travaillées par semaine habituellement, déclarées par le répondant, ne dépassait pas 168 (24 heures x 7 jours) et de demander aux intervieweurs de confirmer le nombre d'heures auprès des répondants lorsque le nombre déclaré d'heures par semaine dépassait 80.

On a procédé à des analyses post-enquête afin d'examiner la fréquence à laquelle les vérifications ont été déclenchées pendant les interviews des répondants. Étant donné qu'aucun de ces chiffres n'était élevé, il était inutile de pousser la vérification de la cohérence interne.

En plus de ces analyses, on a relevé deux cas où on a décelé une erreur dans le questionnaire électronique, ce qui a fait en sorte que des données manquaient pour certains répondants. Dans les deux cas, l'imputation était considérée comme une solution possible. L'imputation consiste à remplacer les données manquantes par une valeur établie en fonction de valeurs provenant d'un répondant donneur présentant des caractéristiques semblables ou qui sont fondées sur des valeurs établies par déterminisme logique.

Dans le premier cas, on a oublié un sous-groupe de répondants (n = 581) dans le module sur les modifications liées à l'emploi à la question EMO_Q02, qui demandait si on avait offert au répondant d'apporter certaines modifications à son lieu de travail. Plus particulièrement, en raison d'une erreur de programmation commise dans l'enchaînement des questions, le sous-groupe de répondants oublié se composait de personnes hors du marché de travail avec incapacité, qui ont travaillé depuis 2007, qui ont dit que leur état ne les empêche pas complètement de travailler et qui auraient besoin d'une accommodation pour travailler. Nous avons considéré l'option d'imputer les réponses manquantes pour ces personnes, en utilisant de l'imputation par donneur. Le problème est que ces personnes ont un profil très particulier et, après avoir analysé de près les donneurs possibles, nous n'avons pas pu trouver des donneurs convenables, ayant des charactéristiques semblables aux répondants. En conséquence, nous avons laissé les réponses manquantes plutôt que d'imputer des réponses provenant de donneurs qui ont des profils (au niveau de l'emploi) très différents.

Dans le deuxième cas, les réponses aux questions SNC_Q01D et SNC_Q01E (« *Dans la dernière année, avez-vous reçu des revenus de prestations d'invalidité du Régime de pensions du Canada ?* » et « *Dans la dernière année, avez-vous reçu des revenus de prestations d'invalidité du Régime de rentes du Québec ?* ») ont été imputées pour certains répondants (n= 466) du fait que leur année de naissance était manquante. En effet, ces questions étaient posées lorsqu'une condition sur l'année de naissance était satisfaite (YOB>=1946). Une année de naissance manquante [7] faisait en sorte que certaines personnes devaient répondre à ces questions mais elles n'étaient pas posées. Comme les deux réponses possibles étaient « oui » ou « non », nous avons imputé leurs réponses à l'aide de deux modèles de régression logistique pour prédire la probabilité de répondre « oui » à ces deux questions, respectivement. Pour construire et valider ces modèles, nous avons utilisé les données des autres répondants avec incapacité qui avaient donné une année de naissance et qui avaient donc bien respecté le cheminement des questions. Nous avons ensuite appliqué les modèles aux cas pour qui les données étaient manquantes afin d'imputer une valeur, selon la probabilité prédite de dire « oui ». Des indicateurs identifiant les cas imputés ont été créés pour les utilisateurs.

En ce qui concerne les vérifications externes, on a procédé à des comparaisons entre les données de l'ECI et celles de l'ENM, pour les indicateurs où les concepts se chevauchaient de façon significative et où des incohérences auraient risqué de présenter des problèmes d'interprétation aux fins d'analyse pour les utilisateurs. Une vingtaine d'indicateurs ont été analysés, essentiellement dans les secteurs de l'activité sur le marché du travail, du niveau de scolarité et de la situation familiale. L'hypothèse de base était que l'ECI contient des données plus récentes que l'ENM et les variables de l'ENM, dans leur version de diffusion, ont possiblement été assujetties à plusieurs manipulations et imputations. La règle générale est donc la suivante : en présence d'une incohérence, la valeur de la variable de l'ENM en cause est effacée et remplacée par une valeur spéciale, « 95 », permettant à l'utilisateur d'en tenir compte durant leurs analyses.

5.8 Variables dérivées et variables externes couplées de l'ENM

Afin de faciliter l'analyse approfondie du vaste ensemble de données de l'ECI, on a créé une centaine de variables dérivées en regroupant ou en combinant des éléments du questionnaire. De plus, environ 200 variables de l'ENM ont été couplées au fichier de traitement final de l'ECI pour 2012.

7 L'application utilisait le champ correspondant à la date de naissance rapportée directement par le répondant au moment de l'entrevue, mais certains répondants ne rapportaient que leur âge.

On a créé des variables dérivées afin de fournir aux utilisateurs des indicateurs de l'état d'incapacité et du type d'incapacité, en fonction des définitions utilisées pour l'ECI (consultez les définitions d'incapacité pour l'enquête à la section 2.1). On a également créé des variables dérivées afin de saisir la sévérité des incapacités et les catégories de sévérité entre les types d'incapacité. On en a également créé pour l'utilisation d'appareils et accessoires fonctionnels et l'aide nécessaire dans les activités quotidiennes, ce qui obligeait à combiner plusieurs éléments dans le questionnaire. De multiples concepts importants liés au marché du travail ont été saisis par des variables dérivées, comme l'ont été les sources de revenus. Finalement, plusieurs variables dérivées reflétaient le codage des variables par rapport aux systèmes de classification type, comme la Classification internationale des maladies et la Classification des programmes d'enseignement.

En créant des variables dérivées pour les répondants, si une question ne recevait pas de réponse (autrement dit, une valeur « Je ne sais pas », « Refus », « Pas déclaré » ou « Saut de question valide »), on attribuait à la variable dérivée la valeur « Pas déclaré ».

La plupart des noms des variables dérivées ont un « D » à la position du premier caractère du nom. Pour toutes les variables couplées de l'ENM, le nom de la variable de l'ENM a été préservé dans la mesure du possible dans la base de données de l'ECI. Certaines exceptions s'appliquaient étant donné que le nom des variables de l'ECI est limité à huit caractères et que celui des variables de l'ENM dépassait parfois huit caractères.

La liste complète des variables dérivées accompagnées des spécifications pour leur dérivation se trouve dans les dictionnaires de données de l'ECI de 2012 (pour en savoir davantage, communiquez avec les Services à la clientèle de Statistique Canada). Les variables dérivées sont classées par thème à l'annexe A avec d'autres indicateurs d'enquête. La liste complète des variables dérivées et des variables de l'ENM, accompagnées de leurs spécifications se trouve dans les dictionnaires de données de l'ECI de 2012.

6. Pondération et création des fichiers de données définitives

6.1 Pondération

Dans une enquête par sondage, chaque répondant ne représente pas que lui- ou elle-même mais aussi d'autres personnes qui n'ont pas été échantillonnées. C'est pourquoi un poids est assigné à chaque répondant de manière à indiquer le nombre de personnes qu'il représente. Pour assurer la cohérence des données et s'assurer que les résultats représentent bien l'ensemble de la population cible et non seulement les personnes échantillonnées, ce poids doit être utilisé pour calculer toutes les estimations.

Le calcul des poids pour l'ECI se fait en plusieurs étapes. La première étape consiste à assigner un poids initial fondé sur le plan de sondage à chaque unité sélectionnée pour l'ECI. Le poids initial est l'inverse de la probabilité d'inclusion. Pour l'ECI, le poids initial est le produit de deux éléments : le poids à l'ENM et le poids de sous-échantillonnage de l'ECI (l'inverse de la fraction d'échantillonnage). Ensuite viennent un certain nombre d'ajustements aux poids pour tenir compte d'exclusions lors de la collecte, de la non-réponse, et pour éviter les poids extrêmes dans les domaines d'estimation. On termine en calant les poids d'enquête sur des totaux estimés de l'ENM et en procédant à certains ajustements pour tenir compte des unités qui étaient dans le champ de l'enquête lors de la sélection en mai 2011, mais hors champ au moment de l'enquête en 2012. Les sous-sections suivantes décrivent les grandes étapes de la pondération.

Calcul des poids initiaux

Il faut calculer les poids initiaux pour l'échantillon OUI et l'échantillon NON. Puisque le plan d'échantillonnage de l'ECI est fondé sur celui de l'ENM, le poids initial correspond au produit du poids final de l'ENM et de l'inverse de à la fraction de sondage de l'ECI[8]. Le poids final de l'ENM tient compte du plan d'échantillonnage de l'ENM, des cas de non-réponse et d'autres corrections. Pour en savoir davantage sur la stratégie de pondération de l'ENM, reportez-vous à la section 4.3 du guide de l'utilisateur de l'ENM.

Le poids initial de l'ECI correspond donc au produit du poids initial ajusté et corrigé de l'ENM par l'inverse de la fraction de sondage à l'ECI. La fraction de sondage à l'ECI est calculée en prenant la taille de l'échantillon sélectionné dans une strate, divisé par le nombre d'unités disponibles sur la base de sondage dans cette strate.

Les ajustements de poids décrits dans les prochaines sous-sections ne concernent que l'échantillon OUI, parce qu'aucune donnée n'a été recueillie pour l'échantillon NON; il est donc inutile de faire des ajustements pour les cas où aucun contact n'a été établi ou de les cas de non-réponse.

Ajustement pour les unités non envoyées à la collecte

L'échantillon qui a été sélectionné pour l'ECI a été gonflé légèrement en prévision de l'exclusion de certaines unités de la collecte pour les raisons suivantes :

- Sélection de plus de 3 personnes dans un même ménage ;
- Sélection de personnes dans des ménages déjà sélectionnés par d'autres enquêtes dans le Nord (Enquête sur la santé dans les collectivités canadiennes, Enquête sur la population active ou Enquête sur les dépenses des ménages);
- Aucun numéro de téléphone disponible pour un ménage (en dehors des TNO) ;
- Pas de nom ni de date de naissance rapporté à l'ENM pour une personne, donc aucune façon d'identifier le bon répondant dans le ménage.

On avait estimé ces pertes lors du calcul de la taille d'échantillon, et sur-échantillonné dans certaines strates pour compenser. On a donc traité les unités exclues de la collecte comme des non-répondants et on a redistribué leurs poids aux autres unités au niveau de la strate.

8 Le poids de l'ENM qui a été utilisé ici est une version corrigée, ce qui évite d'avoir des poids beaucoup trop élevé.

Ajustements pour le non-contact et la non-réponse

La non-réponse à l'ECI se divisait en deux grandes catégories : le non-contact, et la non-réponse étant donné le contact. Ces deux types de non-réponse étaient traités séparément puisqu'ils représentaient deux phénomènes différents. En effet les facteurs qui expliquaient le non-contact étaient davantage reliés aux caractéristiques du ménage et à la mobilité des personnes, tandis que ceux qui expliquaient la non-réponse dans un ménage contacté étaient davantage reliés aux caractéristiques des individus.

Dans un premier temps, les unités envoyées sur le terrain ont été séparées en deux grands groupes : les unités contactées, et celles non contactées. On s'est servis de la régression logistique pour modéliser la probabilité d'être contacté. Les variables explicatives pour ce modèle provenaient soit de la base de l'ENM, soit des paradonnées[9]. Les variables retenues pour le modèle de non-contact étaient : la province, le fait d'être propriétaire ou locataire de son logement, le nombre de personnes dans le ménage, le nombre d'enfants dans le ménage, le nombre de chambres à coucher, le revenu total du ménage, un indicateur de mobilité dans les derniers 5 ans, le groupe de minorité visible, le fait d'être un indien inscrit, la classe de travailleur, le statut de population active, le type d'occupation, la raison de ne pas avoir pu commencer à travailler la semaine avant le recensement, le groupe d'âge, l'état matrimonial, le fait d'accepter que ses données de l'ENM soient rendues publiques dans 92 ans, la connaissance des langues officielles, le statut de la personne au sein de la structure de famille du recensement et le nombre de tentatives de contact.

Grâce à ce modèle on a obtenu, pour chaque unité (contactée ou non), une probabilité d'être contactée. On a ensuite formé des classes de réponse homogènes en groupant les unités avec une probabilité d'être contactées similaires. On s'est servi d'une méthode automatique de formation des classes[10] afin de générer des classes homogènes en termes de probabilité prédite d'être contacté, et comprenant un nombre suffisant d'unités contactées pour éviter de trop grands facteurs d'ajustement des poids. Au total 37 classes ont été formées, et à l'intérieur de chacune, le poids des unités non contactées a été redistribué aux unités contactées.

Ensuite, un ajustement a été fait pour un sous-ensemble des personnes contactées mais non répondantes, soit celles qui ont une incapacité ou une condition de santé ne leur permettant pas de répondre, ou qui ont répondu au module des QII (et selon leurs réponses ont une incapacité), mais pas au reste de l'entrevue de l'ECI. Puisque le nombre de ces cas était minime (environ 700), on a procédé à un ajustement relativement simple au niveau des strates, où le poids des non répondants avec incapacité a été réparti parmi les répondants avec incapacité.

L'étape suivante consistait à ajuster les poids pour tenir compte des autres non répondants (généralement des refus). Ici encore, on s'est servi de la régression logistique pour modéliser la probabilité de répondre étant donné qu'un contact a eu lieu au niveau du ménage. Les variables retenues pour le modèle de non-réponse avec contact étaient : la province, le groupe d'âge, la présence d'enfants dans le ménage, le nombre d'enfants dans la famille de recensement, le nombre de soutiens dans le ménage, le fait de vivre dans un grand centre urbain ou non, le groupe de minorité visible, la classe de travailleur, le groupe d'occupation, le revenu total du ménage, le fait d'accepter que ses données de l'ENM soient rendues publiques dans 92 ans, la connaissance des langues officielles, le nombre de tentatives de contact, le nombre de contacts, le nombre de refus, le nombre de rendez-vous, le nombre moyen de jours entre chaque contact, la date du dernier contact (on a divisé la période de collecte en 4 groupes), le jour de la semaine du dernier contact et l'heure du dernier contact.

Grâce à ce modèle de non-réponse, on a obtenu pour chaque unité (répondante[11] ou non), une probabilité de répondre étant donné le contact. On a ensuite formé des classes de réponse homogènes en groupant les unités avec une probabilité de répondre similaire. La même procédure que celle utilisée pour le modèle de contact a été utilisée et a résulté en la formation de 14 classes. À l'intérieur de chaque classe, le poids des unités non répondantes a été redistribué parmi les unités répondantes.

9 Les paradonnées sont des variables de collecte disponibles pour toutes les unités sélectionnées, comme par exemple le nombre de tentatives de contact, la date, l'heure et la durée de chaque contact, le résultat de chaque contact (refus, prise de rendez-vous, interview complétée, etc.).

10 La procedure PROC FASTCLUS de SAS a été utilisée ici.

11 À noter ici que les unités hors champs ont été exclues de cet ajustement (mais étaient inclues dans l'ajustement pour le non-contact). En effet, puisqu'il y a déjà eu contact avec le ménage pour tous ces cas, on aurait su si la personne sélectionnée était hors champs. On suppose donc ici que tous les non répondants avec contact sont dans le champ de l'enquête et leur poids est réparti parmi les répondants dans le champ de l'enquête seulement.

À noter que les unités hors champs (les décès, les entrées en institutions, les personnes qui vivent maintenant à l'extérieur du pays, etc.) ont été considérées au départ comme des unités répondantes, dans le sens où on a pu parler à quelqu'un de leur ménage qui nous a confirmé leur statut de hors champs. On n'a pas mis leur poids à zéro, on a plutôt conservé leur poids puisqu'elles représentaient des unités de la population de départ (au 10 mai 2011) qui étaient sorties du champ de l'enquête à l'automne 2012. Ces unités sont toutefois exclues du fichier analytique.

Ajustement pour poids extrêmes

Suite aux ajustements pour le non-contact et la non-réponse, on a vérifié la distribution des poids des répondants pour détecter la présence de très grands poids par province puis par domaine d'estimation. En effet, certains facteurs d'ajustement pouvaient avoir généré de très grands poids pour certains individus comparativement aux autres dans certains domaines, et ceci pouvait avoir un effet néfaste sur les estimations et leur variance. On a donc utilisé une méthode appelée « écart-sigma » pour détecter ces poids extrêmes d'abord à l'intérieur de chaque province, puis à l'intérieur de chaque domaine d'estimation. Un exemple d'application de la méthode du sigma-gap est décrit dans Bernier et Nobrega (1998)[12]. La méthode du sigma-gap telle qu'utilisée ici, visait à détecter les grands écarts entre deux poids successifs (et supérieurs à la médiane), lorsque ceux-ci étaient triés en ordre croissant. Lorsqu'on trouvait un écart qu'on jugeait trop grand entre deux poids successifs, le plus grand poids des deux et tous les suivants étaient alors catégorisés comme étant aberrants. Pour juger de la grandeur d'un écart entre deux poids, on la comparait à un certain nombre de fois l'écart-type de la distribution de tous les poids. Pour l'ECI, on a d'abord cherché à détecter les écarts entre les poids supérieurs à trois fois l'écart-type de la distribution à l'intérieur de chaque province. La valeur de trois a été choisie puisque c'est celle qui permettait d'identifier les poids aberrants qu'on aurait nous-mêmes identifiés si on avait appliqué une procédure manuelle. Tous les poids identifiés comme étant aberrants ont été fixés à la plus grande valeur non aberrante de la province. On a ensuite réappliqué la même procédure mais au niveau des domaines d'estimation. La réduction des poids qui a découlé de cette étape a été compensée à l'étape de la post-stratification, tel que décrit plus bas.

Transfuges de domaines d'estimation

Avant de passer à la post-stratification, on a dû tenir compte des cas dont l'âge dérivé à partir de la date de naissance déclarée à l'ECI ne correspondait plus au groupe d'âge auquel le cas était assigné selon la base de sondage. Cette situation était possible puisque la date de naissance sur la base de l'ENM pouvait être manquante et avoir été imputée, ou pouvait avoir été mal déclarée ou mal saisie. On a trouvé un peu moins de 300 cas dans cette situation. Pour ceux-ci, on a comparé leur poids suite à l'ajustement des poids extrêmes de l'étape précédente par rapport à la distribution des poids dans leur nouveau domaine. Lorsque le poids de l'individu tombait dans la fourchette des poids de son nouveau domaine, celui-ci était conservé tel quel. Mais s'il tombait en dehors de la fourchette des poids de son nouveau domaine, alors on le changeait pour mettre la valeur minimum (si le poids était inférieur à la fourchette) ou maximum (si le poids était supérieur à la fourchette) du nouveau domaine. On ajusté ainsi le poids de 35 personnes dans l'échantillon de l'ECI.

12 Bernier, J. et Nobrega, K. (1998). « Outlier detection in asymmetric samples: A comparison of an inter-quartile range method and a variation of a sigma-gap method ». Rencontre annuelle de la Société Statistique du Canada, juin 1998.

Post-stratification et ajustement pour les hors champs

Deux étapes distinctes de post-stratification ont été nécessaires pour obtenir les poids finaux. Ces étapes ont été nécessaires à la fois pour l'échantillon des OUI et celui des NON.

Première post-stratification

La première post-stratification consistait à ajuster les poids des répondants à l'ECI (c'est-à-dire les hors-champs, les répondants avec incapacité et les répondants sans incapacité) de sorte à obtenir les mêmes totaux qu'à l'ENM pour la population des OUI et ce, par province, groupe d'âge, sexe et sévérité. Par sévérité, on entend la sévérité indiquée à l'ENM, soit « souvent » ou « parfois ». Les groupes d'âge sur lesquels on a post-stratifié sont les suivants : 15-19, 20-24, 25-29,...,60-64, 65-79, 75+.

La même post-stratification a été faite aux poids préliminaires de l'échantillon des NON en partant du poids initial calculé plus tôt.

La première post-stratification a donc été faite de façon indépendante pour les OUI et les NON. Les totaux de contrôle ont été calculés à partir de la base de diffusion de l'ENM pour les 15 ans et plus, vivant dans des logements privés, et excluant les réserves.

Deuxième post-stratification

La deuxième post-stratification avait pout but principal d'ajuster les poids de l'échantillon des NON afin d'estimer la part de la population qui était hors champs à l'automne 2012. Néanmoins, cette étape a été appliquée sur l'échantillon des OUI et des NON. Puisque la collecte de l'ECI a eu lieu environ un an et demi après l'ENM, un grand nombre d'unités hors-champs a été observé dans l'échantillon des OUI, représentant environ 235 000 personnes dans la population des OUI. Puisque ces hors champs doivent être exclus des taux d'incapacité, il est important de tenter d'exclure également les hors champs au dénominateur (comprenant la population des OUI et des NON). Autrement, on risquerait de sous-estimer les taux d'incapacité. Comme l'échantillon des NON n'a pas été envoyé à la collecte, on a dû trouver une méthode indirecte pour estimer et exclure les hors champs. La stratégie utilisée pour estimer les hors-champs chez les NON, comme on le verra, dépend du type de hors-champs et de l'information disponible pour le faire.

Dans un premier temps, on a calé les poids des échantillons de OUI et de NON sur les estimations de population[13] produites par la Division de la démographie, ajustées pour le sous-dénombrement net au 10 mai 2011, excluant les moins de 15 ans, les réserves indiennes, les logements collectifs, les membres des Forces canadiennes et les visiteurs au pays. Cette calibration était nécessaire puisque les totaux de décès et d'émigrants qu'on utilisera ensuite pour estimer les hors champs dans la population des NON correspondent à la population totale et non uniquement à la population dénombrée. Les décès représentent plus de 65 % des hors champs observés à l'ECI. L'ajustement pour ce type de hors champs se devait d'être le plus précis possible.

Dans un second temps, on a procédé aux différents ajustements pour exclure de la population des NON les unités qui, à l'automne 2012, auraient été hors champs. Il ne s'agit pas ici d'identifier quelles personnes dans l'échantillon des NON sont des hors champs, mais plutôt de réduire la somme des poids de l'échantillon des NON afin qu'ils correspondent aux totaux de population excluant les hors champs.

Ces ajustements ont été faits séparément pour différents types de hors champs, et à l'aide de différentes sources de données, tel que décrit ci-après.

Le tableau suivant donne les estimations des différents types de hors champs dans la population des OUI.

13 Au moment de procéder à la pondération de l'ECI, la Division de la démographie produisait encore ses estimations de population à partir d'un modèle basé sur le Recensement de 2006. Les données permettant d'ajuster les données du Recensement de 2011 pour tenir compte du sous-dénombrement net n'étaient pas encore disponibles. On a tout de même jugé qu'il valait mieux faire cet ajustement pour éviter de sous-estimer les taux d'incapacité de l'ECI.

Tableau 6.1
Estimations des types de hors champs dans la population des OUI

Type de hors champs	Non pondéré	Pondéré	Pondéré
	nombre		%
Décès	939	153 570	65,40
Émigrants	66	15 870	6,80
Entrées en institutions	349	57 450	24,50
Déménagé sur réserve indienne	4	120	0,05
Autre	62	7 800	3,30
TOTAL	**1 420**	**234 810**	**100,00**

Source : Statistique Canada, Enquête canadienne sur l'incapacité, 2012.

Décès et émigrants (entre le 10 mai 2011 et la collecte de l'ECI)

La Division de la démographie a produit des totaux cumulatifs de décès et d'émigrants entre le 10 mai 2011 et le 15 novembre 2012 (soit le milieu de la collecte de l'ECI) par province et territoire, groupe d'âge et sexe. Ces totaux étaient ajustés pour exclure les moins de 15 ans, les réserves indiennes, les logements collectifs, les membres des Forces canadiennes et les visiteurs au pays. Puisque ces totaux couvrent l'ensemble de la population (population des OUI et des NON), on pouvait soustraire de ce total le nombre de décès estimés pour la population des OUI grâce à l'échantillon de l'ECI (et aux poids obtenus suite à la post-stratification sur les totaux de démographie tel que décrit plus haut) pour déduire le nombre de ceux-ci provenant de la population des NON. On pouvait ensuite ajuster le poids de sondage des NON à la baisse pour reproduire la population toujours vivante et étant toujours au pays à l'automne 2012.

Tableau 6.2
Estimations du nombre de décès et d'émigrants entre le jour du Recensement de 2011 et la collecte de l'ECI

	Estimation de l'ECI (population des OUI)	Estimation de la Div. de la démographie (population totale)	Estimation indirecte (population des NON)
Décès	153 570	338 520	184 950
Émigrants	15 870	70 290	54 420

Sources : Statistiques Canada, Enquête canadienne sur l'incapacité, 2012 et Programme des estimations démographiques.

L'ajustement à la baisse des poids de la population des NON a été fait pour les décès au niveau des provinces croisées par deux groupes d'âge (15-64 et 65+). Pour l'émigration, l'ajustement a été fait par groupe d'âge seulement (15-24, 25-44 et 45+). Il était important que l'ajustement soit fait au niveau le plus précis possible sans toutefois créer des incohérences dans les ajustements à effectuer.

Personnes vivant en institution après le 10 mai 2011

Ce type de hors champ est plus difficile à estimer pour la population des NON. En effet, il n'existe aucune source de données exacte nous permettant d'estimer le nombre de personnes entrées en institutions depuis le dernier recensement. De plus, on sait que la proportion de personnes entrées en institutions parmi la population des OUI devrait être supérieure à celle parmi la population des NON, mais on ne sait pas à quel point celle-ci est plus grande.

On a obtenu des estimations d'entrées en institution de l'Enquête sur la dynamique du travail et du revenu (EDTR, panel 7) pour la population totale de 15 ans et plus (avec les mêmes exclusions de réserves, logements collectifs, etc. que l'ECI, mais excluant aussi les territoires). Ces estimations nous ont été fournies par groupe d'âge (15-64, 65-74,75+) et couvrent une période de un an et demi, ce qui est comparable au délai entre le Recensement et l'ECI. Comme l'EDTR n'a pas d'estimation pour les territoires, il a été impossible d'estimer le nombre d'entrées en institution chez les NON dans les territoires. Une estimation indirecte des entrées en institution chez les NON a été obtenue en soustrayant le nombre d'entrées en institution/ estimé pour la population des OUI à l'ECI du total estimé pour l'ensemble de la population à partir de l'EDTR (comme on l'a fait pour les décès et l'émigration).

Le nombre d'entrées en institution chez les NON pour les deux premiers groupes d'âge (15-24 et 65-74) était négligeable, on a donc seulement procédé à un ajustement pour le dernier groupe d'âge chez les NON. On a donc réduit le poids des NON de 75 ans et plus afin de réduire le total de population d'un peu moins de 26 000 personnes dans les 10 provinces.

Personnes ayant déménagé sur une réserve indienne après le 10 mai 2011

Le nombre de hors champs chez les OUI dû à un déménagement sur réserve étant tellement faible, on n'a pas jugé nécessaire de procéder à un ajustement chez les NON.

Autres hors champs

Pour les autres types de hors champs, soit les membres des Forces canadiennes, les visiteurs au pays (mauvaise classification à l'ENM) et âge non valide, on a dû supposer que la proportion estimée pour la population des OUI aurait été la même pour la population des NON. Cette hypothèse est moins réaliste pour les membres des Forces canadiennes, toutefois cet ajustement pour les hors champs est minime.

Ainsi on a chez les OUI un total de personnes pour ces 3 types de hors champs représentant environ 0.13% de la population des OUI [14]. On a donc supposé que 0.13% des NON étaient de tels hors-champs, et on a diminué les poids des NON proportionnellement à ce qui a été observé chez les OUI dans les 3 groupes d'âges suivants : 15-24, 25-44, 45+.

Enfin il faut bien comprendre ici que l'exclusion des hors champs chez les NON a été fait en ajustant à la baisse le poids de ceux-ci pour compenser ces pertes. Chez les OUI au contraire, on a simplement exclus du fichier analytique les personnes identifiées comme étant hors champs, et dont le poids de sondage nous avait permis d'estimer le nombre de personnes qu'elles représentaient dans la population des OUI au départ.

14 Ici, 0.13% est calculé en prenant 7739 et le divisant par la somme totale des poids de l'ECI après la deuxième post-stratification.

Comptes pondérés de la population

On retrouve au tableau 6.3 des comptes pondérés de la population des OUI couverte par l'enquête, soit les personnes ayant déclaré une limitation d'activités aux questions filtres de l'ENM, après ajustement pour le sous-dénombrement net et exclusion des hors champs en date de l'automne 2012.

Tableau 6.3
Population des OUI de l'ENM couverte par l'ECI, et nombre de personnes avec incapacité, par province et territoire

Province/Territoire	Population couverte	Personnes avec incapacité
	nombre	
Terre-Neuve-et-Labrador	85 130	59 300
Île-du-Prince-Édouard	26 370	18 840
Nouvelle-Écosse	199 490	143 760
Nouveau-Brunswick	145 300	99 450
Québec	956 330	616 740
Ontario	2 399 620	1 651 620
Manitoba	214 640	145 270
Saskatchewan	172 330	116 640
Alberta	573 660	369 190
Colombie Britannique	891 330	546 760
Yukon	6 300	4 070
Territoires du Nord-Ouest	3 900	2 740
Nunavut	2 900	1 540
Canada	**5 677 290**	**3 775 910**

Source : Statistique Canada, Enquête canadienne sur l'incapacité, 2012.

Le tableau suivant donne des comptes pondérés de la population couverte par l'échantillon des NON, après ajustements pour le sous-dénombrement net et exclusion des hors champs en date de l'automne 2012.

Tableau 6.4
Population des NON de l'ENM couverte par l'échantillon des NON, par province et territoire

Province/Territoire	Population couverte
	nombre
Terre-Neuve-et-Labrador	335 850
Île-du-Prince-Édouard	91 070
Nouvelle-Écosse	565 610
Nouveau-Brunswick	461 520
Québec	5 480 590
Ontario	8 328 280
Manitoba	715 010
Saskatchewan	606 820
Alberta	2 371 490
Colombie Britannique	2 811 680
Yukon	22 060
Territoires du Nord-Ouest	29 470
Nunavut	19 450
Canada	**21 838 910**

Source : Statistique Canada, Enquête canadienne sur l'incapacité, 2012.

6.2 Structure et contenu des fichiers

On a créé pour l'ECI deux fichiers de données : un ficher comprenant les personnes avec incapacité, et un autre comprenant les personnes sans incapacité. Selon le type d'analyse qu'on veut faire, on devra soit utiliser le fichier de personnes avec incapacité seulement, soit utiliser les deux fichiers ensemble.

Le fichier de **personnes avec incapacité** comprend les personnes sélectionnées à l'ECI qui, selon la définition de l'incapacité retenue pour l'ECI, sont considérées comme ayant une incapacité. Ce fichier est le plus complet des deux. Il contient à la fois toutes les données de l'ECI, et plusieurs variables provenant de l'ENM. Toute analyse visant exclusivement les personnes avec incapacité peut être faite grâce à ce seul fichier.

Le fichier de **personnes sans incapacité** quant à lui comprend trois groupes de personnes, les deux premiers groupes provenant de l'échantillon des OUI de l'ECI, le troisième de l'échantillon des NON. Les 3 groupes sont les suivants :

a) les personnes interviewées à l'ECI qui ont déclaré n'être limitées que « Rarement » avec « Aucune difficulté » ou « Peu de difficulté ». Ces personnes étaient, aux fins de l'entrevue de l'ECI, considérées comme ayant potentiellement une incapacité et donc ont eu à répondre à toutes les questions de l'ECI. Toutefois elles ont été exclues de la définition finale de l'incapacité.

b) les personnes interviewées à l'ECI et qui ont déclaré n'être « Jamais » limitées (les « faux positifs »). Ces personnes étaient, aux fins de l'entrevue de l'ECI, considérées comme n'ayant pas d'incapacité et n'ont donc pas eu à répondre au reste des questions de l'ECI.

c) les personnes provenant de l'échantillon des NON. Celles-ci n'ont pas déclaré de limitations d'activité à l'ENM, n'ont pas été envoyées à la collecte, et sont automatiquement considérées comme n'ayant pas d'incapacité.

Le fichier de personnes sans incapacité comprend donc un contenu différent, selon le groupe de personnes à qui on a affaire. Ainsi pour les personnes du groupe « a », on a le même contenu que pour les personnes avec incapacité, c'est-à-dire toutes les données de l'ECI et plusieurs variables provenant de l'ENM. Pour les personnes du groupe « b », on n'a que les données du module des QII provenant de l'ECI puisque l'entrevue s'est terminée immédiatement après ce module. On a cependant aussi les variables de l'ENM. Enfin pour les personnes du groupe « c », on n'a que les variables de l'ENM puisqu'on n'a pas fait de collecte pour ces unités.

Le fichier de personnes sans incapacité doit être utilisé conjointement avec le fichier de personnes avec incapacité pour deux types d'analyses : le calcul de taux d'incapacité, puisque le dénominateur doit inclure à la fois les personnes avec et sans incapacité; puis la comparaison des caractéristiques de l'ENM[15] des personnes avec et sans incapacité.

Pour être en mesure de différencier chaque groupe de personnes dans les fichiers, on a créé une variable dérivée appelée CSDPOPFL qui prend la valeur 1 pour les personnes avec incapacité, la valeur 2 pour les personnes sans incapacité du groupe « a », la valeur 3 pour les personnes sans incapacité du groupe « b », et la valeur 4 pour les personnes sans incapacité du groupe « c ».

Pour faciliter l'utilisation de ces deux fichiers, on y retrouvera les mêmes variables. Cependant certaines variables auront des valeurs manquantes dans le fichier des personnes sans incapacité puisqu'on ne possède pas toutes les informations pour certains groupes de personnes. Le tableau suivant résume le contenu disponible de chaque fichier, et de chaque groupe de personnes pour le fichier des personnes avec et sans incapacité.

15 On parle ici des caractéristiques de l'ENM puisque pour la grande majorité des personnes sans incapacité, on ne possède que ces variables (échantillon des NON).

Tableau 6.5
Contenu disponible selon différents groupes de personnes dans les fichiers analytiques de l'ECI

	CSDPOPFL	Variables démographiques	QII	Autre contenu ECI et variables dérivées	Variables ENM	Poids final et poids Bootstrap[1]
Personnes AVEC incapacité	1	√	√	√	√	√
Personnes SANS incapacité (groupe (a))	2	√	√	√	√	√
Personnes SANS incapacité (groupe (b))	3	√	√		√	√
Personnes SANS incapacité (groupe (c))	4	√			√	√

√ contenu disponible
1. Pour une description des poids Bootstrap, voir la section 7.2.
Source: Statistique Canada, Enquête canadienne sur l'incapacité, 2012.

Note sur les périodes de référence

Lorsqu'on vise à calculer des taux d'incapacité ou lorsqu'on cherche à comparer les caractéristiques des personnes avec et sans incapacité, la **date de référence est le 10 mai 2011**. En revanche, lorsqu'on s'intéresse uniquement aux caractéristiques des personnes avec incapacité, on travaillera avec les données recueillies à l'ECI et telles que mesurées à l'automne 2012. La **période de référence sera donc ici l'automne 2012**. Ceci revient à considérer l'ECI comme un type d'enquête longitudinale où on a recueilli une première vague de données en 2011 grâce à l'ENM (population initiale) autant pour la population des OUI que des NON, et une seconde vague de données en 2012 grâce à l'ECI pour le sous-ensemble de personnes avec incapacité.

En d'autres mots, les personnes avec incapacité de l'ECI sont donc des personnes qui, au 10 mai 2011, avaient déclaré une limitation d'activité à l'ENM et qui, en 2012, ont déclaré une incapacité lors de l'ECI. Les caractéristiques de l'ECI pour les personnes avec incapacité représentent donc des **informations de 2012** au sujet d'une **population définie en 2011**.

6.3 Ensembles de données finaux et dictionnaires de données

Les fichiers de données finaux comprennent :

- Fichier de traitement final
- Fichiers analytiques à utiliser dans les Centres de données de recherche.

Le fichier de traitement final est un fichier maison qui comprend un certain nombre de variables temporaires utilisées exclusivement aux fins de traitement. Les fichiers analytiques sont des fichiers de diffusion qui ont fait l'objet d'un traitement plus approfondi aux fins de publication. Ils sont distribués dans les Centres de données de recherche au Canada. À Statistique Canada, ils servent également à produire des tableaux de données à la demande de clients. On a pris de nombreuses mesures afin d'assurer la confidentialité des données des répondants. Les fichiers de diffusion devraient être distribués après le jour de la diffusion des résultats de l'ECI le 3 décembre 2013 (pour obtenir des précisions sur la diffusion, veuillez consulter le chapitre 9).

Afin de transformer le fichier de traitement final et nettoyé en fichiers de données analytiques finaux, une série de mesures ont été effectuées. D'abord, des mesures ont été adoptées afin d'assurer une meilleure protection des renseignements confidentiels fournis par les répondants. Ensuite, les poids affectés à la personne ont été ajoutés aux fichiers. La pondération est décrite de façon plus détaillée à la section 6.1. Enfin, toutes les variables temporaires ou les variables utilisées exclusivement à des fins de traitement ont été supprimées des fichiers.

Voici les documents complémentaires qui accompagnent les fichiers analytiques de l'ECI de 2012 :

- le cliché d'enregistrement,
- la syntaxe liée à SAS (Système d'analyse statistique), à SPSS (Ensemble des programmes statistiques relatif aux sciences sociales) et à Stata, afin de charger les fichiers,
- les métadonnées sous la forme de dictionnaires de données qui décrivent chaque variable et présentent les comptes de fréquence pondérés et non pondérés.

6.4 Lignes directrices pour l'analyse

On a élaboré un guide de l'utilisateur pour les fichiers analytiques. Ce guide décrit étape par étape la façon d'utiliser les fichiers de données de l'ECI. Le guide de l'utilisateur comprent des lignes directrices pour la mise en tableau et l'analyse statistique, de l'information sur la façon d'appliquer les poids nécessaires aux données, de l'information concernant des progiciels disponibles et des lignes directrices pour la diffusion des données, comme les règles d'arrondissement. La procédure permettant d'estimer la fiabilité des estimations quantitatives et qualitatives est décrite de façon détaillée. De plus, on dispose de deux manuels des codes détaillés, qui consistent en dictionnaires de données pour toutes les variables disponibles.

7. Qualité des données

7.1 Aperçu de l'évaluation de la qualité des données

L'Enquête canadienne sur l'incapacité (ECI) vise à produire des estimations de qualité concernant le type et la sévérité des incapacités qu'éprouvent les Canadiens de 15 ans et plus (en date du 10 mai 2011). Elle fournit également des estimations à l'égard de divers indicateurs importants des expériences et des difficultés des personnes ayant une incapacité. Le présent chapitre examine la qualité des données de l'enquête.

Les sections 7.2 et 7.3 exposent deux genres d'erreur qui se produisent dans les enquêtes — les erreurs d'échantillonnage et les erreurs non dues à l'échantillonnage. Chaque type d'erreur est évalué dans le cadre de l'ECI. Une erreur d'échantillonnage correspond à la différence entre les données obtenues à partir de l'échantillon de l'enquête et les données qui auraient été obtenues d'un recensement complet réalisé dans des conditions semblables auprès de l'ensemble de la population. Par conséquent, une erreur d'échantillonnage peut être décrite comme la différence découlant de la variabilité des échantillons. Les erreurs non dues à l'échantillonnage désignent toutes les autres erreurs causées par des facteurs autres que l'échantillonnage. Les erreurs non dues à l'échantillonnage peuvent se produire à n'importe quelle étape du processus d'enquête et elles comprennent l'absence de réponse à une enquête ainsi que les erreurs introduites pendant la collecte des données ou le traitement informatique.

Le présent chapitre décrit les différentes mesures adoptées afin de prévenir les erreurs lorsque c'était possible et d'apporter des ajustements en cas d'erreur lors des différentes étapes de l'ECI. Les aspects qu'il faut traiter avec prudence dans l'interprétation des données de l'ECI sont notés. Les lecteurs peuvent également consulter le guide de l'utilisateur de l'Enquête nationale auprès des ménages pour obtenir des renseignements connexes sur la qualité des données.

7.2 Erreurs d'échantillonnage et méthode Bootstrap

Les estimations qui peuvent être calculées à partir de cette enquête sont fondées sur un échantillon de personnes. Des estimations quelque peu différentes auraient pu être obtenues si l'on avait procédé à un recensement complet au moyen des mêmes questionnaires, intervieweurs, superviseurs, méthodes de traitement, etc. que ceux réellement utilisés. La différence entre une estimation obtenue à partir de l'échantillon et celle découlant d'un dénombrement complet effectué dans des conditions similaires est appelée « erreur due à l'échantillonnage » de l'estimation.

Afin de fournir des estimations de l'erreur due à l'échantillonnage pour les statistiques produites dans le cadre de l'ECI, on a eu recours à un type particulier de méthode « bootstrap ». Plusieurs méthodes bootstrap existent dans les ouvrages publiés, mais aucune d'elles n'était appropriée pour le plan d'échantillonnage complexe de l'ECI. Les particularités du plan d'échantillonnage qui font que l'estimation des erreurs dues à l'échantillonnage est difficile sont les suivantes :

- Plan d'échantillonnage à trois phases dans lequel des ménages (ou logements) sont sélectionnés aux deux premières phases, et des personnes le sont à la troisième. À la première phase, un échantillon aléatoire de 4,5 millions de ménages stratifié par unité de collecte (UC) a été sélectionné pour participer à l'ENM. À la mi-juillet 2011, un sous-échantillon de 2e phase de 400 000 ménages parmi les 1,2 million qui n'avaient pas encore répondu a été sélectionné dans le cadre de l'opération du suivi de la non-réponse (SNR). À la 3e phase, un échantillon d'environ 45 500 individus avec limitations d'activité selon l'ENM a été tiré pour l'ECI.
- Fraction de sondage de l'échantillon de première phase (ENM) non négligeable (environ 1/3 dans les régions N1) et fraction de sondage de l'ECI assez élevée dans certaines strates.
- Strates de l'ECI (combinaisons de domaines d'estimation, régions N1 ou N2, répondant initial vs répondant SNR, souvent vs. parfois limité) non imbriquées dans les strates de l'ENM (UC ou groupes d'UC).
- Méthode utilisée devant être suffisamment souple pour produire des statistiques standards, comme des proportions, des totaux, des moyennes et des ratios, mais aussi des statistiques plus poussées, y compris des centiles, des coefficients de régression logistique, etc.

Pour les besoins du calcul de l'erreur due à l'échantillonnage, on a considéré que les répondants au SNR formaient un échantillon de 3e phase, où la probabilité d'inclusion d'un logement était égale à sa propre probabilité de répondre et ce, de façon indépendante pour chaque logement.

Tel que mentionné précédemment, plusieurs méthodes bootstrap existent dans les ouvrages publiés pour l'échantillonnage à une phase. La plus courante est appelée bootstrap « avec remise » et consiste à sélectionner M sous échantillons avec remise à partir de l'échantillon principal et à produire des estimations pour chaque sous échantillon. L'estimation de la variance bootstrap est par la suite calculée comme fonction du carré des différences entre les estimations provenant de chacun des M sous-échantillons bootstrap et l'estimation provenant de l'échantillon de l'enquête.

L'utilisation de poids bootstrap simplifie grandement le calcul de la variance. Pour chaque sous échantillon (réplique bootstrap), les poids d'échantillonnage initiaux sont rajustés en fonction du sous échantillonnage bootstrap, ce qui produit ce que l'on appelle les « poids bootstrap initiaux ». Étant donné que chaque échantillon bootstrap est le résultat de la sélection d'unités avec remise, une unité peut se retrouver plusieurs fois dans un échantillon bootstrap particulier. On peut démontrer que les poids bootstrap sont une fonction du poids initial de l'observation multiplié par ce que l'on appelle « la multiplicité » de l'unité dans l'échantillon bootstrap, c'est à dire le nombre de fois où l'unité est sélectionnée dans l'échantillon bootstrap. La multiplicité d'une unité dans l'échantillon bootstrap est une variable aléatoire qui suit ce que l'on appelle une « distribution multinomiale ». Ainsi, les poids bootstrap peuvent être considérés comme le produit des poids d'échantillonnage initiaux par un facteur de correction aléatoire (qui est, dans ce cas-ci, une fonction de la multiplicité de l'unité). Une fois les poids bootstrap initiaux calculés, tous les ajustements de poids appliqués aux poids d'échantillonnage initiaux sont appliqués aux poids bootstrap initiaux en vue d'obtenir les poids bootstrap finaux, qui rendront compte de la variance liée non seulement au plan d'échantillonnage particulier, mais aussi à tous les ajustements de poids appliqués à l'ensemble de l'échantillon pour calculer les poids finaux.

En 2006, l'Enquête auprès des peuples autochtones (EAPA) avait élaboré une méthode bootstrap généralisée pour l'échantillonnage à deux phases (Langlet, Beaumont et Lavallée, 2008). Cette méthode n'avait pas pu être utilisée pour l'Enquête sur la participation et les limitations d'activités (EPLA) puisque les travaux étaient toujours en cours au moment de la diffusion des données. Pour 2011, on a donc adapté la méthode à deux phases utilisée par l'EAPA de 2006 pour tenir compte des nouvelles complexités du plan de sondage. Une revue de la méthode de 2006 est décrite plus bas.

Tel que mentionné précédemment, les poids bootstrap peuvent être considérés comme le produit des poids d'échantillonnage initiaux et d'un facteur de correction aléatoire. Il s'agit du concept qui sous-tend la méthode du bootstrap généralisée. Dans le cas d'un échantillon à deux phases, la variance peut être décomposée en deux composantes, chacune étant associée à une phase de l'échantillonnage. La méthode du bootstrap généralisée à deux phases produit un facteur de correction aléatoire pour chaque phase d'échantillonnage. Dance ce cas, le poids bootstrap initial d'une unité est le produit du poids de sondage initial par ces deux facteurs d'ajustement aléatoires.

Or en 2011, pour l'estimation de la variance, on considère que l'ENM a deux phases additionnelles, soit celle correspondant au sous-échantillon du SNR ainsi qu'une phase additionnelle associée à la non-réponse au sous-échantillon du SNR. Pour tenir compte de ces phases supplémentaires, on a combiné les trois phases de l'ENM en une seule phase pour ensuite utiliser la méthode du bootstrap généralisé à deux phases (une phase à l'ENM et une phase à l'ECI). Dans la méthode du bootstrap généralisé pour plans à deux phases, on peut montrer que les facteurs d'ajustement aléatoires dépendent des probabilités d'inclusion simples et doubles associées à chacune des phases. Pour combiner les trois phases en une seule phase, il s'agit de combiner les probabilités d'inclusion simples et doubles pour les trois phases de l'ENM. Les probabilités d'inclusion simples et doubles des trois phases combinées sont données par le produit des probabilités d'inclusion simples et doubles de chacune des trois phases. Les détails de la méthodologie utilisée se trouvent dans Haddou (2013)[16].

16 Haddou, M. (2013). «Bootstrap Variance Estimation Specifications - Aboriginal Peoples Survey». Document interne, janvier 2013.

Une fois les trois phases de l'ENM combinées en une seule phase, on utilise la méthode bootstrap généralisée pour l'échantillonnage à deux phases, qui implique le calcul de deux ensembles de facteurs de correction aléatoires, soit un pour chacune des phases.

La présence de ces deux ensembles facteurs de correction aléatoires comporte un avantage majeur. Le premier ensemble peut être utilisé pour les estimations fondées sur la première phase seulement, c'est-à-dire les estimations fondées sur l'échantillon de l'ENM. Ces estimations sont utilisées lorsque les poids sont ajustés en fonction des totaux de l'ENM au moment de la post-stratification (section 6.1). Cette méthode produira des totaux de l'ENM variables pour chaque échantillon bootstrap. Cela rend compte du fait que les totaux de l'ENM utilisés sont fondés sur un échantillon et non sur des totaux fixes connus.

Dans le cas de l'ECI, 1 000 ensembles de poids bootstrap ont été produits au moyen de la méthode décrite précédemment. La méthode utilisée est légèrement biaisée, en ce sens qu'elle surestime légèrement la variance. L'ampleur de la surestimation est considérée comme négligeable pour l'ECI. La méthode peut aussi produire des poids bootstrap négatifs. Pour remédier à ce problème, on a procédé à une transformation des poids bootstrap en vue de réduire leur variabilité. Par conséquent, la variance calculée à partir de ces poids bootstrap transformés doit être multipliée par un facteur qui est fonction d'un certain paramètre appelé phi. La valeur du paramètre est choisie de façon à correspondre au plus petit nombre entier faisant en sorte que tous les poids bootstrap sont positifs. Pour l'ECI, ce facteur est 4. Les variances calculées à partir des poids bootstrap transformés doivent donc être multipliées par $4^2 = 16$. De même, les coefficients de variation obtenus (racine carrée de la variance divisée par l'estimation proprement dite) doivent être multipliés par 4. Toutefois, la plupart des logiciels qui produisent des estimations de l'erreur d'échantillonnage à partir de poids bootstrap comportent une option permettant de préciser ce facteur de correction, afin que la bonne estimation de la variance soit obtenue sans la nécessité d'une étape additionnelle de multiplication par la constante.

Il est extrêmement important d'utiliser le facteur multiplicatif approprié pour n'importe quelle mesure de l'erreur d'échantillonnage, telle la variance, l'erreur type et le c.v. L'omission de ce facteur multiplicatif produira des résultats et des conclusions erronés. Ce facteur est souvent spécifié comme étant « l'ajustement de Fay » dans les logiciels produisant des estimations de l'erreur d'échantillonnage à partir de poids bootstrap.

Pour des exemples de procédures utilisant l'ajustement de Fay, voir le Guide de l'utilisateur des données de l'Enquête sur l'incapacité de 2012.

Enfin, la mesure de l'erreur due à l'échantillonnage utilisée pour l'ECI est le coefficient de variation (c.v.) de l'estimation, c'est-à-dire l'erreur type de l'estimation divisée par l'estimation proprement dite. Dans le cas de cette enquête, lorsque le c.v. d'une estimation est supérieur à 16,5 % mais inférieur ou égal à 33,3 %, l'estimation sera accompagnée de la lettre « E », ce qui indique que les données doivent être utilisées avec prudence. Lorsque le c.v. d'une estimation est supérieur à 33,3 %, ou si l'estimation est basée sur un échantillon de 10 personnes ou moins, l'estimation de la cellule sera remplacée par la lettre « F », ce qui indique que le chiffre a été supprimé pour des raisons de fiabilité.

7.3 Erreurs non dues à l'échantillonnage

Mis à part l'échantillonnage, plusieurs facteurs, qui peuvent se produire à presque toutes les étapes d'une enquête, peuvent causer des erreurs dans les résultats de l'enquête. Les répondants peuvent ne pas comprendre les questions et y répondre de façon erronée, les réponses peuvent être entrées de façon incorrecte lors de la saisie des données, et des erreurs peuvent être introduites dans le traitement et la mise en tableaux des données. Ce sont tous là des exemples d'erreurs non dues à l'échantillonnage.

Sur un grand nombre d'observations, les erreurs aléatoires auront peu d'effet sur les estimations tirées de l'enquête. Toutefois, les erreurs qui se produisent systématiquement contribueront à des biais dans les estimations de l'enquête. Voilà pourquoi, on a consacré beaucoup de temps et d'efforts à réduire les erreurs non dues à l'échantillonnage dans l'enquête. À l'étape de l'élaboration du contenu, de nombreuses activités ont été entreprises afin de formuler des questions qui seraient bien comprises des répondants. Le nouveau questionnaire a fait l'objet de tests poussés au cours de plusieurs séries d'essais qualitatifs. De plus, beaucoup d'initiatives ont été prises sur le terrain afin d'encourager la participation et de réduire le nombre de non-réponses. Les nombreuses méthodes d'assurance de la qualité appliquées aux étapes de la collecte de données, du codage et du traitement des données afin de vérifier et de corriger les erreurs dans les données sont également importantes. Des corrections de la pondération ont été appliquées afin de corriger le biais potentiel, au besoin.

Les facteurs suivants sont souvent à l'origine des erreurs non dues à l'échantillonnage : absence de réponse, couverture, mesure et traitement des données. Les sections suivantes traitent des différentes mesures utilisées pour réduire et corriger les erreurs pouvant de produire dans chacun de ces domaines.

Erreurs de couverture

Des erreurs de couverture se produisent lorsque la population échantillonnée exclut des personnes qui devraient faire partie de la population ciblée. Étant donné que l'ECI est le prolongement de l'ENM de 2011, elle hérite des problèmes de couverture de cette enquête, qui hérite elle-même des problèmes de couverture du recensement de 2011. Pour en savoir davantage au sujet des erreurs de couverture dans le cadre du recensement, veuillez consulter : « Estimations définitives de la couverture du Recensement de 2011 » sur le site Web de Statistique Canada. Pour de plus amples renseignements sur la qualité des données de l'ENM, veuillez consulter le Guide de l'utilisateur de l'ENM sur le site Web de Statistique Canada.

Erreurs de non-réponse

Il y a erreurs de non-réponse lorsqu'on ne parvient pas à recueillir des renseignements complets sur toutes les unités dans l'échantillon sélectionné. La non-réponse entraîne des erreurs dans les estimations de l'enquête de deux façons. D'abord, les non-répondants présentent souvent des caractéristiques différentes des répondants, ce qui peut introduire des biais dans les estimations de l'enquête si l'erreur de non-réponse n'est pas corrigée adéquatement. Dans ce cas-là, plus le taux de non-réponse est élevé, plus le biais sera important. Ensuite, si le taux de non-réponse est plus élevé que prévu, il réduira la taille réelle de l'échantillon. Par conséquent, la précision des estimations sera réduite (l'erreur d'échantillonnage dans les estimations augmentera). Ce deuxième aspect peut être réglé si on sélectionne dès le départ un plus gros échantillon. Cependant, le biais potentiel dans les estimations ne s'en trouvera pas réduit pour autant.

L'ampleur de la non-réponse varie. D'une part, il existe la non-réponse partielle, selon laquelle le répondant omet de répondre à une ou plusieurs questions, mais remplit une partie importante préétablie de l'ensemble du questionnaire. De façon générale, la portée de la non-réponse partielle est demeurée limitée dans le cadre de l'ECI en raison des examens qualitatifs importants qui ont été effectués et de la mise à l'essai des questions. D'autre part, il y a la non-réponse totale, selon laquelle la personne qui a été choisie pour participer à l'enquête n'a pas pu être jointe ou a refusé de participer à l'enquête après avoir été contactée. Le poids des répondants a été augmenté afin de tenir compte de l'absence de réponse, tel que décrit à la section 6.1.

Dans le but de réduire le nombre de non-réponses, on a fait preuve de nombreuses initiatives avant et pendant la collecte des données.

Le site Web de Statistique Canada comprenait une page Web dédié à l'ECI qui renfermait une série de questions et de réponses pour les répondants, ainsi que des renseignements généraux au sujet de l'enquête. Dans les mois précédant l'enquête, un lien spécial donnant accès au site Web de Statistique Canada a été placé sur le site Web d'organisations à l'intention des personnes ayant une incapacité, afin de leur donner accès à des renseignements utiles concernant l'ECI à venir. Avant la collecte des données, chaque répondant sélectionné a reçu un dépliant accompagné d'une lettre de présentation qui donnait un aperçu de l'enquête et expliquait l'importance de participer. Un petit feuillet en braille a également été distribué.

De plus, les intervieweurs ont reçu une formation poussée. Les intervieweurs ont été formés par un personnel expérimenté de Statistique Canada. Des manuels d'intervieweur détaillés on été fournis comme référence, en conjonction avec la formation. De plus, tous les intervieweurs étaient sous la direction d'intervieweurs principaux, qui supervisaient les activités sur le terrain. Les intervieweurs principaux, qui encourageaient les répondants à participer à l'enquête, n'ont pas ménagé leurs efforts non plus pour joindre les non-répondants en les rappelant et en faisant des appels de suivi. Le tableau des taux de réponse définitifs obtenus pour l'ECI est présenté à la section 4.2 du présent guide. Le taux de réponse globale pour l'enquête s'établissait à 74,6 %. Les taux de réponse étaient les plus élevés dans les groupes de personnes plus âgées.

Erreurs de mesure

Il y a erreurs de mesure lorsque la réponse donnée est différente de la valeur réelle. Ce type d'erreur peut être attribuable au répondant, à l'intervieweur, au questionnaire, à la méthode de collecte des données ou encore au système de conservation des dossiers du répondant. Dans le cadre de l'ECI de 2012, on a mené un travail considérable afin d'élaborer des questions qui seraient comprises, pertinentes et sensibles aux besoins des répondants.

Plusieurs séries d'essais qualitatifs ont été menées dans le cadre de l'ECI, notamment pour les nouvelles questions d'identification des incapacités (QII). Les essais qualitatifs ont été effectués par le Centre de ressources en conception de questionnaires (CRCQ) de Statistique Canada. Afin de réduire les erreurs de mesure, des modifications ont été apportées à la formulation et l'enchaînement des questions.

On a également pris plusieurs autres mesures afin de réduire les erreurs de mesure, notamment le recours à des intervieweurs compétents, en donnant une formation approfondie à ces derniers en matière de méthodes d'enquête et de questionnaires, et l'observation et le suivi du travail des intervieweurs afin de cerner les problèmes liés à la conception du questionnaire ou une mauvaise interprétation des instructions.

Erreurs de traitement

On peut commettre des erreurs de traitement à différentes étapes, notamment la programmation de l'application ITAO, la saisie des données par l'intervieweur, le codage et la vérification. Des procédures de contrôle de la qualité ont été appliquées à chaque étape du traitement des données, afin de réduire ce type d'erreur. Les interviews pour l'ECI ont été réalisées au moyen d'interviews assistées par ordinateur, et plusieurs vérifications ont été effectuées dans le système afin de permettent aux répondants de confirmer les valeurs inhabituelles pendant l'interview, et pour corriger les erreurs immédiatement[17].

À l'étape du traitement des données, on s'est servi d'une série précise de procédures et de règles de vérification afin de repérer et de corriger les incohérences entre les réponses fournies. Pour chaque étape du nettoyage des données, on a mis au point un ensemble de procédures systématiques complètes afin d'évaluer la qualité de chaque variable du fichier et de corriger chaque erreur relevée. Un aperçu des fichiers de sortie a été établi à chaque étape, et on a procédé à une vérification en comparant les fichiers à l'étape courante et à l'étape précédente. La programmation de toutes les règles de vérification a fait l'objet d'essais exhaustifs avant d'être appliquée aux données. À titre d'exemples de vérification du traitement des données, mentionnons l'examen de l'enchaînement des questions, y compris les séquences très complexes, afin de vérifier si les valeurs de sauts de question avaient été correctement affectées et distinguées de différents types de valeurs manquantes, le codage en double aux fins de contrôle de la qualité des réponses « Autre- Précisez », la supervision par du personnel expérimenté du codage par rapport aux classifications types et l'examen des variables dérivées par rapport aux variables de leur composante afin d'assurer la programmation adéquate de la logique de dérivation, y compris des dérivations très complexes. Consultez le chapitre sur le traitement du présent guide pour en savoir davantage.

17 La seule exception à cet égard concerne les interviews sur place effectuées dans les Territoires du Nord-Ouest avec des questionnaires papier dont les données ont été saisies plus tard dans le système ITAO au bureau régional.

8. Différences entre l'Enquête canadienne sur l'incapacité (ECI) de 2012 et l'Enquête sur la participation et les limitations d'activités (EPLA) de 2006

L'Enquête canadienne sur l'incapacité (ECI) de 2012 comporte un ensemble de questions d'identification des incapacités qui étaient utilisées pour la première fois afin d'identifier les personnes avec incapacité au Canada. Certains utilisateurs de données chercheront peut-être à comparer la prévalence de l'incapacité entre les enquêtes, notamment par rapport à l'Enquête sur la participation et les limitations d'activités (EPLA) de 2006 qui a précédé l'ECI. Mais cette comparaison n'est pas possible pour plusieurs raisons.

Les concepts et les méthodes ayant servi à mesurer l'incapacité dans le cadre de l'ECI de 2012 diffèrent considérablement de ceux utilisés pour l'EPLA de 2006. Le principal changement touche la définition de l'incapacité : dans l'ECI, la définition repose sur un nouvel ensemble de « questions d'identification des incapacités » (QII). Ces questions de sélection reflètent une mise en application plus complète du modèle social de l'incapacité, une plus grande cohérence dans l'identification de l'incapacité selon chaque type, de même qu'une meilleure couverture de l'ensemble des types d'incapacité, en particulier les incapacités mentales/psychologiques et cognitives (apprentissage et mémoire)[18]. Ces différences sont décrites de façon plus détaillée dans la suite du présent document.

En raison des différences majeures entre les concepts et les méthodes de l'EPLA de 2006 et de l'ECI de 2012, il n'est pas possible ni recommandé de comparer la prévalence de l'incapacité au fil du temps entre ces deux sources de données.

8.1 Nouvelle méthode d'identification des incapacités

Contrairement aux questions de sélection de l'EPLA qui reposaient sur une approche hybride (modèle social pour l'identification de certains types d'incapacité et modèle médical pour d'autres types), les questions d'identification des incapacités utilisées dans l'ECI (QII) ont été élaborées de manière à assurer une plus grande cohérence de l'identification des incapacités selon le type.

D'après leurs réponses aux questions du module QII, les répondants sont considérés comme ayant une incapacité uniquement si leurs activités quotidiennes sont limitées[19] en raison d'un trouble ou d'une difficulté à accomplir certaines tâches.

L'ECI (puisqu'elle utilise le module QII) permet aux répondants de déterminer s'ils font face à une limitation d'activités en raison de ces difficultés ou de ces troubles. Parfois, des personnes ayant déclaré avoir de la difficulté à accomplir certaines tâches ou avoir un trouble quelconque vont ensuite indiquer que cela ne nuit en rien à leurs activités quotidiennes. Ces personnes étaient considérées comme ayant une incapacité dans le cadre de l'EPLA, mais pas dans le cadre de l'ECI.

Ce changement aura l'incidence la plus marquée sur l'identification des personnes avec une incapacité sensorielle ou physique, puisque dans l'EPLA, l'identification de ces types d'incapacités reposait exclusivement sur l'existence d'un certain degré de difficulté. Cependant, pour les types d'incapacité autres que physiques, l'EPLA présente une plus grande concordance avec l'ECI parce qu'elle comportait elle aussi une exigence additionnelle portant sur les limitations d'activités.

D'autres changements apportés aux questions d'identification peuvent eux aussi avoir un impact sur les résultats. Par exemple, les questions relatives aux incapacités mentales/psychologiques ont été quelque peu modifiées: des exemples des troubles les plus courants (comme la dépression, l'anxiété et le trouble bipolaire) ont été ajoutés et les exemples de troubles moins courants et plus stigmatisés (comme la schizophrénie) ont été supprimés. La liste d'exemples dans les questions relatives aux troubles d'apprentissage et aux troubles de mémoire a aussi été modifiée.

Autre différence entre les deux enquêtes : les troubles de la communication étaient identifiés dans l'EPLA mais ne le sont pas dans l'ECI. Durant la série de tests qualitatifs successifs du module QII, aucune question n'a pu être trouvée afin d'identifier correctement les personnes avec troubles de la communication. La

18 Cet énoncé est vrai en ce qui a trait au module QII lui-même; il existe toutefois certaines contraintes associées à l'administration de ce module dans le contexte de l'ECI de 2012, qui est une enquête postcensitaire. Reportez-vous à la section 8.3 pour en savoir davantage.

19 La seule exception à cet égard concerne les troubles du développement, le répondant qui a reçu un tel diagnostic étant considéré comme ayant une incapacité.

plupart des itérations d'une question servant à identifier les personnes de ce petit groupe (y compris la question utilisée dans l'EPLA) soulevaient des difficultés chez les personnes dont la première langue n'est ni le français ni l'anglais, ou encore des difficultés d'ordre culturel (comme le fait de ne pas comprendre certaines expressions appartenant au langage familier). De plus, l'émergence des médias sociaux comme moyen de communication semble avoir ajouté une nouvelle complexité au concept de communication pour la population canadienne.

Enfin, des changements ont été apportés au concept d'« agilité » utilisé dans l'EPLA. Dans le module QII, l'incapacité correspondante a été subdivisée en deux types (flexibilité et dextérité), les essais qualitatifs ayant démontré que les gens jugent que ces deux aspects sont très différents l'un de l'autre, et qu'ils ne sont pas associés aux mêmes troubles sous-jacents. On a jugé que cette subdivision constituait une amélioration en vue d'identifier différents types d'incapacités physiques; en outre, cela donne suite à des demandes formulées par les utilisateurs de données sur l'incapacité ainsi que par la comité consultatif technique pour les personnes avec incapacités d'Emploi et Développement social Canada (EDSC).

8.2 Autres changements touchant le contenu de l'ECI

De même qu'il n'est pas possible de comparer la prévalence de l'incapacité en général et celle de certains **types** d'incapacité en particulier, l'ECI ne peut pas non plus être comparée à l'EPLA pour ce qui est d'autres données relatives à son contenu. En effet, le contenu de l'ECI a été en grande partie simplifié et mis à jour. Une portion du contenu de l'EPLA a été supprimée en raison de contraintes opérationnelles; toutefois, on s'est efforcé, dans la mesure du possible, de supprimer le contenu s'étant révélé moins pertinent et moins utilisé.

Les questions de l'enquête ont également été mises à jour afin de mieux refléter la réalité d'aujourd'hui et de rectifier certains problèmes connus avec l'EPLA. Par exemple, la section portant sur les aides et appareils fonctionnels a fait l'objet de changements importants. Bon nombre des appareils mentionnés dans l'EPLA étaient considérés comme étant dépassés, compte tenu de l'utilisation qu'en font actuellement les personnes avec incapacité. De même, de nouveaux appareils ont été ajoutés pour rendre compte des progrès technologiques réalisés depuis la fin des années 1990, soit la période où ont été formulées les questions de l'EPLA sur les aides et appareils fonctionnels.

Des efforts ont également été faits afin de simplifier la méthode servant à interroger les répondants de l'ECI au sujet de l'aide requise et des besoins non satisfaits. Par exemple, les questions sur le besoin et l'utilisation d'accessoires (tels que des barres d'appui, etc.) installés dans le logement des personnes avec incapacité physique sont maintenant combinées aux questions sur d'autres aides et appareils associés au même type d'incapacité. Dans l'EPLA, on distinguait les questions sur les aides et appareils portables et portés par la personne elle-même, de celles ayant trait aux aides et appareils installés dans le logement. Cette distinction n'était pas naturelle pour de nombreuses personnes qui voulaient indiquer leur besoin pour des items tels que des barres d'appui, dans la première section relative aux aides et appareils fonctionnels (souvent déclarés dans la catégorie « autres »), puis qui en faisaient de nouveau état dans une section subséquente lorsqu'on leur posait précisément la question, pouvant donner lieu à un double compte.

En raison de ces changements apportés au libellé et à l'enchaînement des questions, il convient d'éviter les comparaisons entre les données de l'EPLA et celles de l'ECI.

8.3 Questions filtres de l'ENM

Bien que l'on considère que l'ECI et les questions d'identification des incapacités (QII) constituent un progrès important en vue de mieux mesurer l'incapacité au moyen du modèle social, il convient de mentionner que l'échantillon de l'ECI a été présélectionné à l'aide des mêmes questions filtres à l'Enquête nationale auprès des ménages (ENM) de 2011, que celles utilisées dans le questionnaire complet du Recensement de 2006 lors de l'EPLA. Des études de suivi ont montré que ces questions filtres ne permettent pas d'identifier adéquatement les personnes ayant une incapacité mentale/psychologique ou cognitive. Cela signifie que l'ECI présente encore quelques-unes des lacunes que l'on retrouvait dans l'EPLA en ce qui touche le sous-dénombrement de certains types d'incapacité. Néanmoins, parmi les personnes sélectionnées sur la base de l'ENM, la nouvelle méthode de sélection utilisée pour l'ECI contribuera à améliorer l'identification des personnes ayant une incapacité mentale/psychologique, une incapacité cognitive ou un « autre » type d'incapacité, car il est maintenant possible de mieux les identifier.

8.4 Changement relatif au délai suite à l'enquête de présélection

Comme mentionné précédemment, les bases de sondage de l'EPLA et de l'ECI ont été établies à partir des réponses au questionnaire complet du Recensement de 2006 et à l'ENM de 2011, respectivement. L'EPLA a été menée de six à neuf mois après le Recensement, tandis que l'ECI a été menée sur le terrain de 16 à 20 mois après l'ENM. En raison du délai plus long, non seulement a-t-il été plus difficile de faire le suivi des répondants sélectionnés qui avaient déménagé, mais cela a en outre fait augmenter la possibilité qu'un répondant ayant déclaré une limitation d'activité à l'ENM n'ait plus d'incapacité, qu'il ait été institutionnalisé ou qu'il soit décédé dans l'intervalle (voir la section 8.5). Il se peut aussi que certains renseignements recueillis dans le cadre de l'ENM et annexés au fichier de l'ECI (par exemple, des données sur le revenu) aient changé au cours de la période séparant les deux enquêtes.

8.5 Autres changements d'ordre méthodologique

La base de sondage de l'ECI a été constituée à partir des réponses à l'ENM de 2011, tandis que celle de l'EPLA reposait sur les réponses au questionnaire complet du Recensement de 2006. Bien que tous les efforts aient été déployés par l'ENM afin de réduire les effets du taux de réponse plus faible, ce changement particulier a pu avoir une incidence sur les résultats de l'ECI. Pour en savoir davantage sur le suivi des cas de non-réponse et la correction de la pondération pour l'ENM, veuillez vous reporter au chapitre 6.

Une deuxième différence méthodologique entre l'ECI et l'EPLA découle d'un changement apporté à la stratégie de pondération dans le cadre de l'ECI pour contrebalancer le délai plus long entre la collecte de l'ENM et celle de l'ECI. Comme mentionné plus tôt, ce délai entraînait une augmentation de la probabilité de non-réponse à l'ECI en raison du décès ou de l'institutionnalisation de certains répondants. En raison de la possibilité que bon nombre de ces personnes soient des personnes avec incapacité, il était important de prendre des mesures afin de ne pas sous-estimer la prévalence de l'incapacité. Par conséquent, on a ajusté les poids applicables à la population ayant répondu NON aux questions filtres de l'ENM de manière à prendre en compte les décès et les institutionnalisations qui pourraient avoir eu lieu entre la collecte des données de l'ENM et celle des données de l'ECI. Cela a nécessité un calage des poids aux estimations de population ajustées pour le sous-dénombrement net, ce qui n'avait pas été fait dans le cas de l'EPLA de 2006. Pour de plus amples renseignements sur la pondération employée dans la cadre de l'ECI, veuillez vous reporter au chapitre 6.

8.6 Résumé et recommandation

Tel que discuté précédemment, les principales différences entre l'EPLA et l'ECI peuvent être resumées ainsi:

- La définition de l'incapacité utilisée pour l'ECI est différente de celle utilisée pour l'EPLA. L'ECI a adopté le nouvel ensemble de questions d'identification des incapacités (QII), qui est utilisé pour la première fois afin d'identifier les incapacités au Canada.
- Les questions de sélection utilisées dans l'ECI concordent davantage avec un modèle social de l'incapacité que celles de l'EPLA; de plus, contrairement à ces dernières, les questions de sélection de l'ECI sont plus cohérentes pour identifier l'incapacité selon chaque type.
- Le contenu du questionnaire a été simplifié et mis à jour afin de tenir compte des technologies actuelles et de corriger certaines lacunes dans la formulation des questions.
- Le délai plus long entre l'ENM et l'ECI a fait augmenter la possibilité que des répondants sélectionnés n'aient plus d'incapacité, soient institutionnalisés ou soient décédé au cours de l'intervalle entre les deux enquêtes. Il a donc été nécessaire de recourir à une différente méthode de calage des poids, ce que l'on n'a pas fait dans le cas de l'EPLA de 2006.
- Enfin, la base de sondage de l'ECI a été constituée à partir de l'ENM de 2011, plutôt que du Recensement de 2006. Ce changement peut avoir eu une incidence sur les résultats de l'ECI.

Il faut présumer que ces différents changements ont une incidence sur la comparabilité des données. Par conséquent, il n'est pas possible ni recommandé de faire des comparaisons entre les données de l'ECI et celles de l'EPLA.

9. Diffusion des données

9.1 Produits et services de données

Les données de l'Enquête canadienne sur l'incapacité (ECI) de 2012 ont été diffusées le 3 décembre 2013. La diffusion comprenait un court feuillet d'information concernant l'incapacité au Canada et une série de tableaux de données sur les taux d'incapacité pour les adultes au Canada, selon l'âge et le sexe, pour chaque province et territoire. Les tableaux comprenaient également des données sur les types d'incapacité et leur sévérité. Le grand public peut consulter ces produits gratuitement sur le site Web de Statistique Canada.

À partir de 2014, les chercheurs au pays pourront procéder à des analyses approfondies en se servant des fichiers de données analytiques de l'ECI disponibles dans les centres de données de recherche (CDR) de Statistique Canada. Pour accéder aux fichiers, les chercheurs doivent soumettre leur proposition à un comité de recherche et d'éthique pour approbation. Leur utilisation des données doit se faire conformément aux politiques, aux directives et aux normes de Statistique Canada. Par exemple, seules des estimations statistiques agrégées conformes aux dispositions de confidentialité de la *Loi sur la statistique* peuvent être diffusées à l'extérieur de Statistique Canada.

La prochaine diffusion comprendra un rapport analytique traitant de sujets importants comme l'utilisation d'aides et d'appareils fonctionnels, l'aide reçue ou requise, et l'expérience de travail et les expériences scolaires des personnes ayant une incapacité.

En plus de ces produits et services de données, les clients peuvent adresser une demande de tableaux de données personnalisés à Statistique Canada. Toutes ces demandes sont filtrées pour s'assurer qu'elles sont conformes aux normes de confidentialité et les données agrégées sont arrondies avant d'être divulguées aux clients. Statistique Canada fait également des présentations spéciales de l'ECI aux principaux intervenants et lors de différentes conférences.

9.2 Produits de référence

L'information sur l'Enquête canadienne sur l'incapacité (ECI) de 2012 peut être consultée sur le site Web de Statistique Canada. Statistique Canada offre une base de métadonnées intégrée (BMDI) en ligne de toutes les enquêtes qu'elle mène, y compris l'ECI de 2012. La BMDI a pour objectif de fournir de l'information qui aidera le public à interpréter les données publiées par Statistique Canada. Cette information (également désignée sous le nom de métadonnées) est publiée pour que les gens comprennent les concepts de base qui définissent les données, y compris les variables et les classifications, les méthodes statistiques et les enquêtes sous-jacentes, de même que les principaux aspects de la qualité des données. Les gens ont également accès directement au questionnaire.

En plus de la BMDI, un guide de l'utilisateur a été rédigé pour les fichiers de données. Dans le but de renseigner le grand public, le guide des concepts et des méthodes de l'ECI est disponible en ligne et fourni des renseignements pertinents sur le contenu, le plan de sondage, les méthodes de collecte, le traitement des données, la qualité des données, les différences entre l'ECI de 2012 et l'EPLA de 2006, de même que les produits de diffusion pour l'ECI.

Les chercheurs qui utilisent les fichiers analytiques des Centres de données de recherche (CDR) de Statistique Canada peuvent consulter un guide de l'utilisateur des CDR, qui décrit, étape par étape la façon d'utiser le fichier de données. Le guide de l'utilisateur des CDR décrit en détails la structure du fichier de données, notamment les variables de base, les variables dérivées et les recoupements avec l'Enquête nationale auprès des ménages (ENM). Les dictionnaires des données comprennent toutes les variables disponibles. Le guide de l'utilisateur des CDR présente également des lignes directrices précises concernant la mise en tableaux et l'analyse statistique, la façon d'appliquer les poids nécessaires aux données, l'information sur les progiciels disponibles et les lignes directrices concernant la diffusion des données, comme les règles d'arrondissement. La procédure consistant à évaluer la fiabilité des estimations, quantitatives et qualitatives, est traitée en détail.

9.3 Contrôle de la divulgation

Il est légalement interdit à Statistique Canada de mettre à la disposition du public toute donnée susceptible de révéler des renseignements obtenus en vertu de la *Loi sur la statistique* qui se rattachent à un particulier, une entreprise ou une organisation identifiable sans que cette personne, entreprise ou organisation donne, par écrit, son consentement à leur révélation. Diverses règles de confidentialité s'appliquent à toutes les données étant diffusées ou publiées afin d'empêcher la publication ou la divulgation de toute information jugée confidentielle. Au besoin, des données sont supprimées pour empêcher la divulgation directe ou par recoupement de données reconnaissables.

Annexe A – Indicateurs du sondage selon le module du questionnaire

La liste ci-dessous contient la description des indicateurs du questionnaire ainsi que les variables dérivées qui combinent différents éléments du questionnaire.

Pour consulter la liste complète des variables, reportez-vous au dictionnaire des données de l'ECI de 2012. Pour de plus amples renseignements sur la façon d'obtenir les dictionnaires des données, veuillez communiquer avec les Services à la clientèle de Statistique Canada ou en composant le 1-800-263-1136.

Questions d'identification des incapacités (QII)

Les QII évaluent le type et la sévérité des incapacités en posant des questions sur la fréquence à laquelle les activités quotidiennes sont limitées par un état à long terme, des problèmes de santé et des difficultés dans les activités.

Questions filtres sur les incapacités

Il y a six questions filtres dans les QII. Les cinq premières questions consistent à connaître la fréquence des difficultés ou des problèmes qui durent depuis six mois ou qui pourraient durer six mois ou plus. Elles portent sur cinq « familles » de catégories d'incapacité :

1. Incapacités sensorielles : difficulté à voir ou à entendre
2. Incapacités physiques (mobilité, flexibilité, dextérité, douleur) : difficulté à marcher, à utiliser des escaliers, à se servir de ses mains ou de ses doigts ou à faire d'autres activités physiques
3. Incapacités cognitives (troubles d'apprentissage, de la mémoire ou du développement) : difficulté à apprendre, à retenir de l'information ou à se concentrer
4. Santé mentale ou psychologique : problèmes d'ordre émotionnel, psychologique ou de santé mentale
5. Autres problèmes ou états de santé chroniques

La sixième question filtre dans les QII consiste à savoir à quelle fréquence la difficulté limite les activités quotidiennes (ou, s'il y a plus d'une difficulté, on cherche à savoir quel est le problème de santé le plus limitant).

Les questions filtres dans les QII ont été élaborées afin d'exclure les personnes qui sont très peu susceptibles d'avoir une incapacité. Dans le contexte de l'Enquête canadienne sur l'incapacité (ECI) de 2012, on avait déjà procédé à un filtrage initial au moyen des questions sur la limitation des activités de l'Enquête nationale auprès des ménages (ENM) de 2011. Voilà pourquoi les questions filtres dans les QII, bien qu'incluses dans l'ECI de 2012, n'ont pas été utilisées pour les besoins du filtrage.

Questions de sélection liées aux incapacités

Les questions de sélection des QII évaluent la présence et la sévérité de dix catégories distinctes d'incapacités liées à des difficultés ou à des problèmes de santé qui durent depuis six mois ou qui pourraient durer six mois ou plus. Ces questions consistent à connaître la fréquence à laquelle une difficulté limite les activités quotidiennes pour les types d'incapacité suivants :

1. Vision
2. Ouïe
3. Mobilité
4. Flexibilité
5. Dextérité
6. Douleur
7. Apprentissage
8. Trouble de développement
9. Troubles mentaux ou psychologiques
10. Mémoire

Les variables dérivées de l'identification des incapacités comprennent l'état de l'incapacité ainsi que les limitations relatives à la vue, à l'ouïe, à la mobilité, à l'agilité, à la flexibilité, à la dextérité, à l'apprentissage, à la douleur, au développement, à l'incapacité mentale ou psychologique, à la mémoire et au type de limitation inconnu, et le dénombrement des types d'incapacité.

Les variables dérivées connexes comprennent une catégorie de sévérité concernant la vue, l'ouïe, la mobilité, la flexibilité, la dextérité, la douleur, l'apprentissage, l'incapacité mentale ou psychologique, la mémoire, le développement et les incapacités de type inconnu, un indicateur de sévérité global et une catégorie de sévérité globale.

Questions sur les incapacités

Cause principale

Ce module sert à déterminer le principal problème de santé, soit celui qui cause le plus de difficulté ou de limitation d'activité au répondant. Les répondants peuvent donner jusqu'à deux problèmes de santé qui contribuent aux difficultés ou aux limitations d'activités subies par le répondant. L'âge du répondant lorsqu'il a commencé à avoir des difficultés ou des limitations d'activités est aussi demandé.

Parmi les autres variables, mentionnons la Classification internationale des maladies (CIM) pour les conditions principales.

Aides et appareils fonctionnels

Pour chaque limitation d'activité déterminée dans les questions d'identification des incapacités (QII), le répondant doit répondre à des questions au sujet d'aides et d'appareils fonctionnels et spécialisés qu'il utilise et dont il a besoin. Les répondants qui déclarent avoir besoin d'aides, d'appareils fonctionnels ou d'autres types d'équipement spécialisé qu'ils n'ont pas doivent répondre à des questions portant sur les raisons pour lesquelles ils n'ont pas ce qu'il faut pour répondre à leur besoin.

Certaines questions portent également sur les dépenses directes liées à l'achat, à la réparation ou à l'entretien des aides, des appareils fonctionnels et de l'équipement spécialisé dont s'est servi le répondant au cours des 12 mois ayant précédé l'enquête.

Les variables dérivées comprennent l'utilisation d'appareils auditifs, le besoin d'appareils auditifs, l'utilisation d'appareils par les personnes ayant une déficience visuelle, le besoin d'appareils pour les personnes ayant une déficience visuelle, l'utilisation d'appareils par les personnes avec un problème de mobilité ou d'agilité, le besoin d'appareils pour les personnes avec un problème de mobilité ou d'agilité, l'utilisation d'appareils par les personnes avec des difficultés d'apprentissage, le besoin d'appareils pour les personnes avec des difficultés d'apprentissage, l'utilisation d'autres appareils par les personnes ayant une autre condition, le besoin d'appareils pour les personnes ayant une autre condition, l'utilisation d'appareils par les personnes avec n'importe quelle condition et le besoin d'appareils pour les personnes avec n'importe quelle condition.

Consommation de médicaments

Ce module porte sur les médicaments sous ordonnance que prend le répondant en raison de son état. On pose également au répondant des questions portant sur l'incidence des coûts des médicaments sur leur consommation.

Aide reçue pour les activités de tous les jours

Ce module porte sur le type d'aide que peuvent recevoir les répondants pour effectuer les activités quotidiennes en raison de leur état. Les répondants doivent inclure l'aide reçue de la part de la famille, d'amis, de voisins et d'organismes, que ce soit à titre gracieux ou contre rémunération. Les activités quotidiennes comprennent la préparation des repas, les travaux ménagers quotidiens, les gros travaux ménagers, les déplacements pour aller à des rendez-vous et pour faire des courses, la gestion des finances personnelles, les soins personnels, l'aide pour des soins médicaux de base à domicile, la mobilité à la maison et la garde d'enfant.

On demande aux répondants qui ne reçoivent pas d'aide normalement s'ils croient qu'ils ont besoin d'aide pour chacune des activités. Les répondants qui déclarent recevoir de l'aide pour plus d'une activité doivent répondre à des questions supplémentaires au sujet de la fréquence de l'aide reçue et au sujet des personnes ou des organismes qui les aident dans leurs activités quotidiennes.

Les répondants doivent également répondre à des questions portant sur les dépenses directes pour l'aide reçue dans les activités quotidiennes au cours des 12 derniers mois.

Veuillez noter que cette section a trait uniquement à l'aide « humaine » et non à l'assistance technique.

Les variables dérivées comprennent le besoin d'aide pour ce qui est de préparer les repas, faire le ménage, faire les gros travaux ménagers, se rendre à des rendez-vous, gérer ses finances personnelles, pour les soins personnels, pour les soins médicaux de base, pour se déplacer et pour s'occuper des enfants, et le fait de recevoir de l'aide.

Éducation

Fréquentation scolaire récente

Ce module porte sur la fréquentation scolaire récente, le statut à temps plein ou à temps partiel, et le type d'établissement institutionnel fréquenté. Les répondants qui fréquentent actuellement une école ou qui ont fréquenté l'école à tout moment depuis septembre 2011 doivent répondre à des questions supplémentaires au sujet de leurs besoins en matière d'installations modifiées, de transports adaptés, d'appareils ou de services de soutien. Si un besoin a été nommé, les répondants doivent dire si les mesures d'adaptation avaient été mises à leur disposition..

Fréquentation passée de l'école

Ce module porte sur la fréquentation passée de l'école et sur le fait que le répondant avait ou non cet état au moment de fréquenter l'école. Les répondants ayant un problème de santé qui fréquentaient l'école à quelque moment que ce soit depuis septembre 2007 se font poser des questions supplémentaires au sujet de leurs besoins en matière d'installations modifiées, de transport adapté, d'appareils ou de services de soutien. Si un besoin a été nommé, les répondants devaient dire si les mesures d'adaptation avaient été prises.

Expériences pédagogiques

L'objet de ce module est de déterminer la portée de l'état de santé d'un répondant sur son expérience pédagogique et le niveau de scolarité. Les répondants doivent répondre à des questions portant sur les difficultés qu'ils ont éprouvées pour étudier. Les questions comprennent les répercussions d'avoir un problème de santé sur les choix de cours ou de carrière, le genre d'école ou le lieu, les dépenses supplémentaires et la période de temps nécessaires pour terminer les études. De plus, les questions portant sur les répercussions sociales perçues de l'état de santé en ce qui a trait à l'exclusion et à l'intimidation sont posées.

Antécédents scolaires

Ce module porte sur le plus haut niveau de scolarité terminé, l'endroit où les études ont été suivies et le principal domaine d'études du plus haut certificat, diplôme ou grade obtenu.

Parmi les autres variables liées à l'éducation, mentionnons le principal domaine d'études (CPE Canada 2011).

Emploi

État de population active

Ce premier module de la section sur l'emploi sert à déterminer, parmi les trois situations d'emploi – personnes ayant un emploi, chômeur, inactifs (comprend les personnes à la retraite) – laquelle s'applique au répondant.

Si un répondant travaillait la semaine précédant l'entrevue, il est classé comme occupant un emploi. S'il était absent du travail uniquement à cause d'une maladie, d'une maladie dans la famille, de vacances, d'un congé de maternité ou de paternité ou d'autres raisons, mais qu'il occupait toujours un emploi, il est classé comme occupant un emploi.

Si un répondant dit qu'il est absent du travail à cause d'une mise à pied temporaire, d'une mise à pied saisonnière, d'une blessure ou d'un congé de maladie qui n'est plus couvert par l'employeur, il est classé comme étant au chômage au moment de l'entrevue. Si le répondant a pris des dispositions pour commencer un emploi au cours des quatre semaines suivant le moment de l'entrevue, il est classé comme étant au chômage.

Un répondant entre dans la catégorie inactif s'il n'a pas d'emploi et qu'il ne peut pas commencer un emploi dans les quatre semaines à venir en raison de ses études, d'une maladie, d'une incapacité ou d'obligations personnelles et familiales, ou s'il ne cherche pas d'emploi.

La retraite est d'abord classée dans la catégorie des inactifs, et passera à la catégorie « à la retraite » dans les modules suivants.

Détails liés à l'emploi

Les répondants doivent répondre à des questions au sujet de leur emploi actuel, y compris les heures travaillées, la situation d'emploi (employé ou travailleur autonome), les détails au sujet de leur entreprise (nom de l'entreprise, genre d'entreprise, industrie, nombre d'employés sur les lieux de travail, etc.) et la profession.

Les répondants doivent également répondre aux questions portant sur les effets de leur état de santé sur leur travail. On demande aux répondants si le fait d'avoir un problème de santé limite la quantité ou le type de travail qu'ils peuvent faire et si leur état a un effet sur leur capacité de changer d'emploi ou d'obtenir de l'avancement dans leur emploi actuel.

Personnes sans emploi

Ce module comporte des questions servant à déterminer le moment où le répondant a travaillé pour la dernière fois et des questions qui portent sur des emplois précédents, y compris le nombre d'heures travaillées, la situation d'emploi (employé ou travailleur autonome), les détails au sujet de l'entreprise (nom de l'entreprise, genre d'entreprise, industrie, etc.) et la profession. On demande aux répondants si le fait d'avoir un problème de santé limite la quantité ou le genre de travail qu'ils peuvent faire et si leur état influe sur leur capacité à chercher du travail.

Hors population active

Les questions de ce module déterminent le moment où le répondant a travaillé pour la dernière fois. Un répondant entre dans la catégorie inactif s'il n'a pas d'emploi et qu'il ne peut pas commencer un emploi dans les quatre semaines à venir en raison de ses études, d'une maladie, d'une incapacité ou d'obligations personnelles et familiales, ou s'il ne cherche pas d'emploi.

Les répondants qui n'ont jamais travaillé doivent répondre à des questions de suivi pour déterminer si leur état de santé les empêche de travailler. Les répondants qui sont retraités de façon permanente doivent dire si cela est dû à leur état de santé.

Les répondants ayant déclaré avoir déjà travaillé par le passé doivent répondre à des questions portant sur leur ancien emploi, y compris les heures travaillées, la situation d'emploi (employé ou travailleur autonome), les détails au sujet de leur entreprise (nom de l'entreprise, genre d'entreprise, industrie, etc.) et la profession.

Les répondants doivent dire si leur état de santé les empêche de travailler ou limite la quantité ou le type de travail qu'ils peuvent faire. On demande également aux répondants des questions portant sur les obstacles auxquels ils doivent faire face au moment de chercher un emploi.

C'est dans ce module que nous distinguons les personnes inactives de celles qui sont à la retraite.

Retraite

Les répondants doivent dire en quelle année ils ont pris leur retraite pour la première fois, si la retraite était volontaire et si leur état de santé les empêche de travailler, ou limite la quantité ou le type de travail qu'ils peuvent faire. Les répondants déclarant une retraite non volontaire doivent répondre à des questions de suivi portant sur les obstacles auxquels ils doivent faire face au moment de chercher emploi.

Formation sur lieu de travail

Les questions de ce module portent sur la formation fournie ou payée par l'employeur ou la compagnie du répondant au cours des 12 mois précédant l'enquête. Une question portant sur les obstacles auxquels ils doivent faire face au moment de suivre une formation liée au travail est également posée.

La formation liée au travail comprend les activités qui sont menées en vue de perfectionner ses compétences ou se recycler pour s'en servir dans un emploi présent ou futur plutôt que par intérêt personnel ou d'autres raisons non liées à l'emploi. Les activités de formation peuvent comprendre les cours et les ateliers ne menant pas à des attestations d'études officielles, ou les activités (programmes) offertes dans les systèmes d'éducation officiels qui mènent à des attestations d'études officielles, y compris les études primaires, secondaires et postsecondaires comme les diplômes et grades universitaires et collégiaux.

Modifications à l'emploi

Les questions de ce module portent sur les modifications ou les changements à l'emploi ou à l'environnement de travail dont le répondant a ou aurait besoin pour travailler. Des questions de suivi servent à déterminer si la modification ou le changement a été apporté, et, si non, pourquoi les mesures d'adaptation n'ont pas été prises.

Discrimination de la population active

Ces questions portent sur la perception qu'a le répondant d'être désavantagé dans la population active, par exemple, se voir refuser une entrevue, des emplois ou des promotions.

Les variables dérivées liées à l'emploi comprennent la situation par rapport au marché du travail, le type d'industrie où la personne travaille (Système de classification des industries de l'Amérique du Nord (SCIAN) Canada 2007), le type de profession (Classification nationale des professions (NOC) 2011), l'emploi à temps partiel ou à temps plein, l'année de l'entrée en fonction dans l'emploi ou l'entreprise actuelle, l'année du dernier emploi et l'année de la première retraite.

Déplacements dans la ville ou communauté

Ce module porte sur les déplacements locaux au moyen de services de transport en commun et adaptés. L'objectif de ce module est de déterminer les difficultés que peut éprouver le répondant pour utiliser ces moyens de transport, par exemple, parce qu'ils ne sont pas disponibles ou pas nécessaires.

Source de revenu

Ce module porte sur le revenu personnel de diverses sources, y compris les salaires et traitements, les prestations d'incapacité, l'assurance-emploi et l'assurance sociale.

Une variable dérivée pour la source de revenus résume les sources de revenus de travail et hors travail.

Annexe B – Indicateur de sévérité des incapacités

Score de sévérité

Un score de sévérité a été développé à partir des questions du module des QII. On décrit ici la méthode utilisée pour dériver un score pour chaque type d'incapacité, puis un score global tenant compte de tous les types d'incapacité.

Le score développé se devait de répondre aux 3 critères suivants :

- le score doit augmenter avec le nombre de types d'incapacité
- le score doit augmenter avec le niveau de difficulté relié à l'incapacité
- le score doit augmenter avec la fréquence de la limitation des activités

Score par type d'incapacité

À partir d'une grille de pointage qui tient compte à la fois de la fréquence des limitations d'activité (Jamais, Rarement, Parfois, Souvent ou Toujours) et de l'intensité des difficultés (Aucune difficulté, Un peu de difficulté, Beaucoup de difficulté ou Incapable de le faire), un score est assigné à chacun des 10 types d'incapacité. Ce score augmente avec la fréquence de la limitation et avec le niveau de difficulté éprouvé. Ainsi une personne déclarant être toujours limitée dans ses activités et être incapable de faire une activité aura le score maximum pour ce type d'incapacité. Voici la grille de pointage qui a été établie pour chaque type d'incapacité. Cette grille donne une plus grande importance aux réponses les plus sévères, une importance moyenne aux réponses modérées, et une plus petite importance aux réponses légères.

Tableau B.1
Grille de pointage basée sur l'intensité de la difficulté et la fréquence des limitations d'activité

Intensité de la difficulté	Fréquence des limitations d'activité				
	Jamais	Rarement	Parfois	Souvent	Toujours
Aucune difficulté	0	0	2[1]	3[1]	4[2]
Un peu de difficulté	0	0	4[2]	6[2]	12[3]
Beaucoup de difficulté	0	3[1]	6[2]	18[3]	20[3]
Incapable	0	4[2]	12[3]	20[3]	24[3]

1. Réponses légères
2. Réponses modérées
3. Réponses sévères

Source : Statistique Canada, Enquête canadienne sur l'incapacité, 2012.

Le score de chacun des types d'incapacité est par la suite normalisé afin de le ramener à une valeur comprise entre 0 et 1. Pour ce faire, le score de chaque cellule est divisé par le score maximum possible, soit 24. Le tableau suivant donne le score normalisé pour chaque type d'incapacité.

Tableau B.2
Pointage normalisé basé sur l'intensité de la difficulté et la fréquence des limitations d'activité

Intensité de la difficulté	Fréquence des limitations d'activité				
	Jamais	Rarement	Parfois	Souvent	Toujours
Aucune difficulté	0	0	0,0833[1]	0,1250[1]	0,1667[2]
Un peu de difficulté	0	0	0,1667[2]	0,2500[2]	0,5000[3]
Beaucoup de difficulté	0	0,1250[1]	0,2500[2]	0,7500[3]	0,8333[3]
Incapable	0	0,1667[2]	0,5000[3]	0,8333[3]	1,0000[3]

1. Réponses légères
2. Réponses modérées
3. Réponses sévères

Source : Statistique Canada, Enquête canadienne sur l'incapacité, 2012.

Cas spéciaux

La grille de pointage donnée précédemment a été utilisée telle quelle pour la majorité des types d'incapacité, mais a due être adaptée pour 3 situations particulières : le type inconnu, cas spécial concerne les troubles du développement, et les types « Mobilité » et « Flexibilité »

D'abord pour le type « Inconnu », les QII ne comprenaient pas de question sur l'intensité de la difficulté. Ainsi on considérait comme ayant une incapacité seulement les personnes ayant déclaré être limitées « Parfois », « Souvent » ou « Toujours ». Puisque par définition les personnes de type « Inconnu » n'ont qu'un seul type d'incapacité, on a examiné la distribution de l'intensité chez les autres personnes n'ayant qu'un seul des 10 types d'incapacité. Il s'avère que dans la grande majorité des cas, le niveau d'intensité était « Un peu de difficulté ». On a donc, pour le type « Inconnu », utilisé la grille de pointage donnée plus haut en fixant le niveau d'intensité à « Un peu de difficulté ». Ainsi les réponses « Parfois » recevaient 0.1667 points, les « Souvent » 0.25 et les Toujours 0.50 points.

Le second cas spécial concerne les troubles du développement. Pour ce type d'incapacité, on considère que dès qu'une personne déclare avoir été diagnostiqué avec ce trouble, elle a une incapacité et ce, peu importe le niveau de difficulté et la fréquence des limitations déclarés. Toutefois lorsque vient le temps de calculer un score pour ce type, on veut éviter d'avoir des personnes avec un trouble du développement qui auraient un score nul ou très petit. On a ainsi fixé un score minimum pour les personnes avec troubles du développement. On utilise ainsi la grille pour déterminer le score, mais si la fréquence de la limitation déclarée est « Jamais », ou si elle est « Rarement » combinée à « Aucune difficulté » ou « Peu de difficulté », alors on fixe le score à la valeur de 0.0833, qui correspond au score minimum qu'une personne peut avoir pour n'importe quel autre type d'incapacité. On a fixé un score minimum pour environ 1 cas sur 5 ayant un trouble du développement.

Le troisième cas spécial concerne les types « Mobilité » et « Flexibilité ». Pour ces deux types, on retrouve non pas une, mais deux questions portant sur l'intensité de la difficulté. Pour calculer le score pour ces deux types, on a utilisé l'intensité maximum déclarée aux deux questions, combinée à la fréquence de la limitation pour décider du score. La décision d'utiliser le maximum des deux questions plutôt qu'une moyenne par exemple, est motivée par le fait qu'on a cherché à donner la même importance à chacun des 10 types dans le calcul du score global. Le fait qu'on ait besoin de 2 questions pour bien établir si le répondant à de la difficulté avec sa mobilité ou sa flexibilité ne devrait pas diminuer le score de cette personne lorsqu'une des deux questions indique « Aucune difficulté » tandis que l'autre indique « Incapable de le faire », par exemple.

Classe de sévérité par type d'incapacité

Pour **chaque type d'incapacité**, on définit 2 classes de sévérité : les moins sévères et les plus sévères. Les personnes ayant un score de 0.5 ou supérieur pour un type d'incapacité sont classées parmi les plus sévères, les autres sont classées parmi les moins sévères pour ce type. Les personnes avec un score de 0 pour un type d'incapacité sont classées comme n'ayant pas ce type d'incapacité.

On a créé une variable de classe de sévérité pour chaque type d'incapacité. (**DVIS_CL** concernant la vue, **DHEA_CL** concernant l'ouïe, **DMOB_CL** concernant la mobilité, **DFLE_CL** concernant la flexibilité, **DDEX_CL** concernant la dextérité, **DPAI_CL** concernant la douleur, **DLEA_CL** concernant l'apprentissage, **DDEV_CL** concernant le développement, **DEMO_CL** concernant l'incapacité mentale ou psychologique, **DMEM_CL** concernant la mémoire, et **DUNK_CL** concernant les incapacités de type inconnu). Chaque variable prend les valeurs suivantes :

0 = aucune sévérité

1 = moins sévère

2 = plus sévère

Score global

Le **score global** de sévérité pour une personne est obtenu en faisant la moyenne des scores des 10 types d'incapacité. Ainsi plus le nombre de type d'incapacité est grand, plus le score sera élevé. Par définition, le score global est également compris entre 0 et 1. Les personnes n'ayant pas déclaré d'incapacité pour les 10 types mais ayant déclaré une autre condition ou un problème de santé (donc de type « inconnu ») ont un score global correspondant au score du type « Inconnu » divisé par 10.

Classe de sévérité globale

Afin de faciliter l'utilisation du score de sévérité, des classes de sévérité ont été établies. Il est important de comprendre que le nom assigné à chaque classe a pour seul but d'en faciliter l'utilisation. Il ne constitue pas une forme d'étiquette ou de jugement du niveau d'incapacité de la personne. En d'autres termes, la façon d'interpréter ces classes est de dire que les personnes de la classe 1 ont une incapacité moins sévère que celles de la classe 2, ces dernières ayant une incapacité moins sévère que celles de la classe 3, et ainsi de suite.

1 = Incapacité légère

2 = Incapacité modérée

3 = Incapacité sévère

4 = Incapacité très sévère

Le point de coupure entre les classes 2 et 3 a d'abord été déterminé de façon graphique. En effet il existe une coupure naturelle dans la distribution du score global autour de la valeur de 0.1. Cette valeur correspond en fait au score d'une personne qui aurait obtenu le nombre de points maximum pour un seul type d'incapacité. Cette personne serait donc dans la classe des sévères. Pour qu'une personne soit classée très sévère, il lui faut le double de ce score, soit 0.2. Un score de 0.2 correspond par exemple à une personne qui aurait obtenu le nombre maximum de points à deux types d'incapacité. Enfin, le point de coupure entre la classe 1 et la classe 2 correspond à la moitié du score requis pour faire partie de la classe 3, soit 0.05. Cela correspond par exemple au score d'une personne qui aurait un score dans une des cellules les plus sévères pour un type d'incapacité.

Le tableau suivant donne la répartition des personnes avec incapacité selon les 4 classes de sévérité.

Tableau B.3
Distribution des personnes avec incapacité selon la classe de sévérité, Canada

Classe de sévérité	Personnes avec incapacité	
	nombre	%
Classe 1 = léger	1 195 590	31,7
Classe 2 = modéré	747 980	19,8
Classe 3 = sévère	849 540	22,5
Classe 4 = très sévère	982 810	26,0

Source : Statistique Canada, Enquête canadienne sur l'incapacité, 2012.

Score global pour les personnes sans incapacité à l'ECI

Dans l'échantillon de l'ECI (l'échantillon des OUI), on retrouve deux types de personnes sans incapacité :

a) les personnes interviewées à l'ECI qui ont déclaré n'être limitées que « Rarement » avec « Aucune difficulté » ou « Peu de difficulté ». Ces personnes étaient, aux fins de l'entrevue de l'ECI, considérées comme ayant potentiellement une incapacité et donc ont eu à répondre à toutes les questions de l'ECI. Toutefois elles ont été exclues de la définition finale de l'incapacité.

b) les personnes interviewées à l'ECI et qui ont déclaré n'être « Jamais » limitées (les « faux positifs »). Ces personnes étaient, aux fins de l'entrevue de l'ECI, considérées comme n'ayant pas d'incapacité et n'ont donc pas eu à répondre au reste des questions de l'ECI.

On a calculé un score global pour les personnes sans incapacité du groupe « a », mais celles du groupe « b » ont toutes un score de 0. Des données de l'ECI étant disponibles pour les personnes du groupe « a », l'utilisateur pourrait vouloir étudier leurs caractéristiques et tenter de comparer le niveau de sévérité des membres du groupe entre eux, bien qu'ils soient tous très légers. Le pointage utilisé pour calculer le score de ces personnes a été intentionnellement fixé à de très petites valeurs pour éviter que celles qui ont déclaré être « Rarement » limitées avec « Aucune difficulté » ou « Peu de difficulté » à plusieurs types finissent avec un score plus élevé qu'une personne considérée comme ayant une incapacité selon la définition retenue pour l'ECI.

Tableau B.4
Pointage utilisé pour les personnes sans incapacité du groupe (a)

Intensité de la difficulté	Fréquence des limitations d'activité				
	Jamais	Rarement	Parfois	Souvent	Toujours
Aucune difficulté	-	0,1	-	-	-
Un peu de difficulté	-	0,2	-	-	-
Beaucoup de difficulté	-	-	-	-	-
Incapable	-	-	-	-	-

Source : Statistique Canada, Enquête canadienne sur l'incapacité, 2012.

Encore une fois, le score de chacun des types d'incapacité est par la suite normalisé afin de le ramener à une valeur comprise entre 0 et 1, donc en divisant celui-ci par le score maximum possible de 24. Le tableau suivant donne le score normalisé des personnes considérées sans incapacité mais ayant répondu à toutes les questions de l'ECI.

Tableau B.5
Pointage normalisé utilisé pour les personnes sans incapacité du groupe (a)

Intensité de la difficulté	Fréquence des limitations d'activité				
	Jamais	Rarement	Parfois	Souvent	Toujours
Aucune difficulté	-	0,0042	-	-	-
Un peu de difficulté	-	0,0083	-	-	-
Beaucoup de difficulté	-	-	-	-	-
Incapable	-	-	-	-	-

Source : Statistique Canada, Enquête canadienne sur l'incapacité, 2012.

Il faut noter toutefois que pour les personnes sans incapacité du groupe « a », bien qu'on ait eu à calculer un score pour chaque type d'incapacité afin d'obtenir le score global, le score de chaque type a par la suite été remis à 0 dans le fichier final. On a également mis le nombre de types d'incapacité à 0. Le but ici est d'éviter toute confusion lorsqu'on étudie les types d'incapacité, ces personnes ne devant pas compter parmi les personnes ayant une incapacité.

Annexe C – Questions sur l'équité en matière d'emploi

Personnes ayant une incapacité à long terme

Questions filtres : DSQ_Q01-Q06

Désavantage dans le domaine de l'emploi

EDE_Q15, UDE_Q12, NDE_Q13 : Est-ce que votre état limite la quantité ou le genre de travail que vous pouvez faire dans votre emploi actuel?

EDI_Q04 : Vous considérez-vous désavantagé sur le plan du travail à cause de votre état?

Perception

EDE_Q19 : Croyez-vous qu'à cause de votre état il vous serait difficile de changer d'emploi ou d'obtenir de l'avancement dans votre emploi actuel?

EDI_Q05 : Croyez-vous que votre employeur actuel ou qu'un employeur éventuel vous considérerait comme étant désavantagé sur le plan du travail à cause de votre état?

Aménagement

EMO_Q01 : À cause de votre état, avez-vous besoin d'une ou de plusieurs des choses suivantes pour être capable de travailler?

1. Une redéfinition des tâches (tâches modifiées ou différentes)
2. Le télétravail
3. Un horaire flexible ou une réduction du temps de travail
4. Un soutien humain, comme un lecteur, un interprète en langage gestuel, un instructeur ou un assistant personnel
5. Une aide technique, comme un synthétiseur vocal, un ATS, un système à infrarouge ou un appareil portable de prise de notes
6. Un ordinateur ou un ordinateur portatif avec un logiciel spécialisé ou autres mesures d'adaptation telles que le braille, un logiciel de grossissement de texte, un logiciel de reconnaissance vocale ou un scanneur
7. Des services de communication, comme la conversion de l'imprimé en braille, en gros caractères ou de l'équipement d'enregistrement
8. Un poste de travail modifié ou ergonomique
9. Une chaise spéciale/un soutien dorsal
10. Des mains courantes, des rampes d'accès ou des portes d'entrée ou des couloirs élargis
11. Un stationnement adapté ou accessible
12. Des ascenseurs adaptés
13. Des toilettes adaptées
14. Du transport adapté
15. D'autres éléments d'équipement, d'aide ou d'aménagement spécial

Annexe D – Catégories supplémentaires de codage et classifications types utilisées

Catégories de réponses supplémentaires créées pour les questions « Autre – Précisez »

Le tableau ci-dessous résume les catégories supplémentaires ajoutées à certaines questions de l'enquête pendant le codage des données, à partir des questions concernant une catégorie « Autre – Précisez ». Lorsque le nombre le justifiait, on a créé des catégories afin de saisir les thèmes dans les données qui n'étaient pas reflétés dans les catégories existantes. Ces catégories sont donc des ajouts aux catégories qui apparaissent déjà dans le questionnaire de l'Enquête canadienne sur l'incapacité et elles pourront être consultées dans la base de données de l'enquête.

Tableau D.1
Catégories de réponse supplémentaires établies pour questions "Autre – précisez"

Thème	Indicateur	Questions	Catégories supplémentaire
Aides et appareils fonctionnels	Les aides ou les appareils fonctionnels dont besoin, mais n'ont pas	AAD_Q10	-Lunettes et/ou lentilles de contact
Aides et appareils fonctionnels	Les aides ou les appareils fonctionnels dont besoin, mais n'ont pas	AAD_Q27	-Aides destinées aux personnes sourdes ou malentendantes -Aides à la mobilité telles que chaises roulantes, triporteurs, cannes et marchettes

Source : Statistique Canada, Enquête canadienne sur l'incapacité, 2012.

Classifications types employées pour l'ECI 2012

Voici la liste des classifications types utilisées dans le codage des questions à développement dans l'Enquête canadienne sur l'incapacité (ECI).

Problème médical : Classification internationale des maladies, dixième révision (CIM-10)

Les répondants à l'ECI ont donné des renseignements sur le principal problème médical qui leur causait le plus de difficulté ou les limitait le plus dans leurs activités quotidiennes. Les réponses ont été codées à partir de la Classification internationale des maladies (CIM-10).

La CIM est un outil de diagnostic standard en épidémiologie, en gestion de la santé et en clinique. Grâce à lui, on peut analyser l'état de santé général de groupes de population. La CIM sert à surveiller l'incidence et la prévalence des maladies ainsi que d'autres problèmes de santé.

La CMI-10 a reçu l'appui de la 43e Assemblée mondiale de la santé en mai 1990 et elle est utilisée par les États membres de l'Organisation mondiale de la Santé (OMS) depuis 1994.

Principal domaine d'études : Classification des programmes d'enseignement (CPE)

Les répondants à l'ECI 2012 ont donné des renseignements sur le principal domaine d'études du plus haut certificat, diplôme ou grade obtenu. Les réponses ont été codées à partir de la Classification des programmes d'enseignement (CPE) Canada 2011.

La CPE a été lancé en 1980 par le National Center for Education Statistics (NCES) aux États-Unis. Le NCES a publié des mises à jour en 1985, 1990, 2000 et 2010. La CPE Canada 2011 est la seconde version canadienne de cette classification, la première version canadienne étant la CPE Canada 2000.

La CPE est conçue pour classer les « programmes d'enseignement ». La CPE s'organise autour du principe du « domaine d'études ». À Statistique Canada, le domaine d'études fait référence à « la discipline ou domaine d'apprentissage ou de formation ».

Industrie et profession : le Système de classification des industries de l'Amérique du Nord et la Classification nationale des professions

Les répondants à l'ECI ont donné des renseignements sur leur occupation et l'industrie dans laquelle ils travaillaient pendant la semaine de référence de l'enquête. Ces réponses ont été codées à l'aide du Système de classification des industries de l'Amérique du Nord (SCIAN) Canada et de la Classification nationale des professions (CNP). Les données ont été codées à un degré de précision à quatre chiffres et par rapport aux versions à jour ou aux nouvelles versions ainsi qu'aux versions antérieures de ces systèmes de classification comme suit :

SCIAN Canada 2012 - Système de classification des industries de l'Amérique du Nord (SCIAN) Canada

SCIAN Canada 2007 Système de classification des industries de l'Amérique du Nord (SCIAN) Canada

CNP – 2011 : Classification nationale des professions (CNP)

CNP-S – 2006 : Classification nationale des professions pour statistiques (CNP-S)

Le SCIAN est un système de classification des industries qui a été conçu par les organismes statistiques du Canada, du Mexique et des États-Unis. Il a été créé dans la perspective de l'Accord de libre-échange nord-américain. La révision du SCIAN pour 2012 avait pour principal objectif de modifier ou de créer des industries qui reflètent les activités et les technologies nouvelles, émergentes ou en évolution. Le SCIAN est axé sur l'offre ou les principes orientés vers la production (les critères utilisés regroupaient les établissements par industrie dans un système s'apparentant aux structures d'intrants, aux compétences professionnelles et aux processus de production).

La CNP 2011 est la taxonomie officiellement reconnue à l'échelle nationale des professions qui forment le marché du travail canadien et elle en constitue le cadre organisationnel. Elle met à jour la Classification nationale des professions d'Emploi et Développement social Canada et la Classification nationale des professions pour statistiques (CNP-S) de Statistique Canada. Comme dans le passé, cette révision est basée sur des recherches, des analyses et des consultations approfondies des professions menées au pays. Les professions sont définies et regroupées principalement en fonction du travail habituellement exécuté, comme en font foi les tâches, les fonctions et les responsabilités de la profession.

Annexe E – Glossaire des termes de l'enquête

A

Agglomération de recensement
Voir **région métropolitaine de recensement (RMR)**

B

Base de données, ensemble de données
Liste organisée et triée de faits ou de données concernant un ensemble de personnes, de ménages, d'entreprises ou d'autres unités pertinentes. Un ensemble de données de Statistique Canada est généralement produit à partir de données d'enquête ou de données administratives, conservé dans un ordinateur et organisé pour pouvoir être accessible facilement dans le cadre d'une vaste gamme d'applications statistiques.

Base de sondage
Liste, carte ou spécification conceptuelle des unités constituant la population d'enquête, à partir de laquelle des personnes peuvent être sélectionnées. Il s'agit, par exemple, d'un répertoire téléphonique ou du répertoire d'une ville, ou encore d'une liste de membres d'une association ou d'un groupe particulier.

C

Centres de données de recherche (CDR)
Le programme des centres de données de recherche (CDR) fournit aux chercheurs l'accès, dans un environnement sécuritaire géré par Statistique Canada, aux microdonnées des enquêtes sur la population et les ménages. Le programme des CDR fait partie d'une initiative de Statistique Canada, du Conseil de recherches en sciences humaines (CRSH) et d'un consortium d'universités, en vue de renforcer les capacités dans le domaine de la recherche sociale au Canada et d'appuyer la collectivité des chercheurs stratégiques. Le programme est aussi soutenu par la Fondation canadienne pour l'innovation (FCI) et les Instituts canadiens de recherche en santé (ICRS).

Centre de population
Le terme « centre de population » remplace le terme « région urbaine » (tel qu'utilisé au recensement de la population jusqu'en 2006). Un centre de population contient une concentration démographique d'au moins 1 000 habitants et une densité de population de 400 habitants ou plus au kilomètre carré. Toutes les régions situées à l'extérieur des centres de population sont classées dans la catégorie des régions rurales.

Cohorte
Groupe de personnes dans le domaine de la démographie ayant une caractéristique commune, par exemple, toutes les personnes d'une population donnée qui sont nées en 1940, ou toutes les personnes qui ont un problème de santé particulier.

Coefficient de variation (C.V.)
Lors d'une enquête sur échantillon, les résultats de l'échantillon sont utilisés pour estimer ce que les résultats auraient été si l'ensemble de la population avait été mesuré. Dans ce processus d'estimation, un certain niveau d'erreur est inévitable. Le coefficient de variation (C.V.) est une façon d'exprimer l'**erreur d'échantillonnage** liée à une **estimation**. Tout d'abord, on calcule une **erreur type** ou erreur « moyenne » pour l'estimation. Le C.V.est obtenu en divisant l'erreur type de l'estimation par l'estimation proprement dite, la fraction qui en résulte étant exprimée en pourcentage. Plus le C.V. est faible, plus la qualité des données est élevée (voir **Marge d'erreur**).

Confidentialité
Résultat d'un rapport de confiance implicite entre la personne qui fournit les données et la personne ou l'organisation qui les recueille. Ce rapport repose sur l'assurance que les données ne seront pas divulguées sans la permission de la personne. En vertu de la ***Loi sur la statistique***, les données qui permettent d'identifier une personne, une entreprise ou une institution ne peuvent pas être divulguées sans que ces personnes, entreprises ou institutions le sachent ou y consentent.

Contrôle
Processus qui permet de veiller à ce que les données d'enquête soient exactes, complètes et uniformes. Un ensemble de règles ou de conditions de contrôle s'applique à un **ensemble de données**. Les données qui ne respectent pas les conditions sont examinées et corrigées au besoin.

Couverture
Il s'agit de la mesure dans laquelle chaque personne ou unité devant être incluse dans une enquête ou un recensement est dénombrée dans les faits et dénombrée seulement une fois. Des erreurs de couverture se produisent lorsque des personnes ou des unités de l'enquête ou du recensement sont laissées de côté (sousdénombrement) ou dénombrées plus d'une fois (surdénombrement). Des études sont souvent menées par

Statistique Canada pour produire des estimations du sousdénombrement et du surdénombrement d'une enquête ou d'un recensement donné, ou pour examiner des questions connexes. Par exemple, Statistique Canada a étudié et analysé la mesure dans laquelle l'utilisation du téléphone cellulaire a des répercussions sur la couverture dans le cas des enquêtes par téléphone.

D

Diffusion
Processus visant à fournir des produits et des services statistiques au grand public et à des utilisateurs de données particuliers. Statistique Canada diffuse les données et les analyses sous forme de résultats d'enquête, de rapports de recherche, de documents techniques, de périodiques, de produits du recensement et de recueils de recherches. Les produits en ligne vont de 1996 à aujourd'hui. Des données historiques peuvent être retrouvées au moyen du catalogue de la bibliothèque. Les données de Statistique Canada sont aussi distribuées par l'entremise d'un réseau approuvé de bibliothèques de dépôt. L'objectif des activités de diffusion est de produire des données pertinentes au moment opportun, selon des formats utiles et des modes accessibles. Les activités en place à l'appui de la diffusion des produits comprennent les services de consultation des clients, le marketing, les promotions, la formation des utilisateurs et d'autres services à la clientèle.

Distribution de fréquences
Graphique ou tableau montrant combien de fois chaque valeur ou fourchette de valeurs d'une **variable** se retrouve dans un **ensemble de données**. On appelle parfois cela un tableau de fréquences unidimensionnel, afin d'indiquer que la distribution comprend des chiffres pour une variable seulement.

Données
Observations et mesures recueillies dans le cadre d'une enquête, d'un recensement ou d'une autre étude. Faits ou chiffres à partir desquels on peut tirer des conclusions.

Données confidentielles
Il s'agit d'un terme utilisé par Statistique Canada pour décrire les données qui sont assujetties aux dispositions relatives au secret de la ***Loi sur la statistique***. Les données sont considérées comme confidentielles parce qu'elles permettent d'identifier directement une **unité répondante**, par exemple, à partir de son nom,

ou parce qu'elles pourraient permettre l'identification d'unités répondantes particulières, même après la suppression des identificateurs, en raison des détails qu'elles comportent, de leur structure géographique ou de leur format.

E

Écart type
Mesure de l'écart ou de la dispersion d'un ensemble de données par rapport à la moyenne. Il s'agit de la mesure de dispersion la plus largement utilisée. Au niveau mathématique, l'écart type correspond à la racine carrée de la **variance**.

Échantillonnage
Processus de sélection d'une partie de la population à observer, afin d'estimer un élément d'intérêt concernant l'ensemble de la population. Parmi les exemples des différentes méthodes d'échantillonnage figurent : l'échantillonnage aléatoire simple, l'échantillonnage aléatoire stratifié, l'échantillonnage en grappes, l'échantillonnage à plusieurs phases, et l'échantillonnage à plusieurs degrés.

Échantillonnage stratifié, stratification
Procédure d'échantillonnage dans laquelle la population est divisée en sous-groupes ou strates homogènes, la sélection des échantillons se faisant de façon indépendante dans chaque strate.

Enquête nationale auprès des ménages
L'Enquête nationale auprès des ménages (ENM) a été instaurée en remplacement du questionnaire complet du recensement, plus largement connu comme la formule « 2B/2D » du recensement. L'ENM vise à recueillir des données sociales et économiques sur la population canadienne. L'objectif de l'ENM est de fournir des données à de petits niveaux géographiques et pour de petits groupes de population.

Enquête postcensitaire
Enquête dans laquelle toutes les **unités enquêtées** sont sélectionnées à partir de leurs réponses au **Recensement de la population ou à l'Enquête nationale auprès des ménages (ENM)**. Ces enquêtes sont généralement menées peu après le recensement ou l'ENM.

Enregistrement
Données pour une **unité répondante** individuelle, dans un fichier comprenant des données pour toutes les unités répondantes d'une enquête.

Ensemble de données, base de données
Liste organisée et triée de faits ou de données concernant un ensemble de personnes, de ménages, d'entreprises ou d'autres unités pertinentes. Un ensemble de données de Statistique Canada est généralement produit à partir de données d'enquête ou de données administratives, conservé dans un ordinateur et organisé pour pouvoir être accessible facilement dans le cadre d'une vaste gamme d'applications statistiques.

Erreur type
Dans une enquête sur échantillon, les résultats de l'échantillon sont utilisés pour estimer ce que les résultats auraient été si l'ensemble de la population avait été mesuré. L'**erreur d'échantillonnage** désigne le fait que les résultats de l'échantillon pondéré diffèrent dans une certaine mesure des résultats qui auraient été obtenus à partir de la population totale. La différence est désignée comme l'erreur d'échantillonnage. L'erreur d'échantillonnage réelle est évidemment inconnue, mais il est possible de calculer une valeur « moyenne », appelée « erreur type ».

Erreurs
Dans une enquête sur échantillon, les résultats de l'échantillon servent à estimer ce que les résultats auraient été si l'ensemble de la population avait été mesuré. L'exactitude d'une telle **estimation** représente une mesure de la façon dont l'estimation diffère du chiffre correct ou « véritable ». Les écarts par rapport aux chiffres véritables sont appelés « erreurs ». Elles peuvent provenir de nombreuses sources, mais peuvent être regroupées en quelques grandes catégories : erreurs de couverture, erreurs dues à la non-réponse, erreurs de réponse, erreurs de traitement et erreurs d'échantillonnage.

Erreurs de couverture
Les erreurs de couverture ont trait aux personnes ou aux unités laissées de côté par l'enquête (sousdénombrement) ou dénombrées plus d'une fois (surdénombrement).

Erreurs dues à la non-réponse
Les erreurs dues à la non-réponse se produisent lorsqu'il est impossible d'obtenir des réponses complètes d'une personne, d'un ménage ou d'une organisation. Même si certains ajustements pour tenir compte des données manquantes peuvent être effectués pendant le traitement, la non-réponse signifie inévitablement une perte d'exactitude.

Erreurs de réponse
Les erreurs de réponse montrent qu'une réponse peut ne pas être entièrement précise. Le répondant peut avoir mal interprété la question ou peut ne pas connaître la réponse, par exemple, si elle concerne un membre du ménage absent.

Erreurs de traitement
Les erreurs de traitement comprennent les erreurs commises pendant l'entrée des données, le codage, la totalisation ou d'autres formes de traitement des données.

Erreur d'échantillonnage
L'erreur d'échantillonnage rend compte du fait que les résultats de l'échantillon pondéré diffèrent dans une certaine mesure des résultats qui auraient été obtenus auprès de la population totale. La différence est appelée erreur d'échantillonnage. L'erreur d'échantillonnage réelle est évidemment inconnue, mais il est possible de calculer une valeur « moyenne », appelée « **erreur type** ».

Estimation
Utilisation des résultats de l'échantillon pondéré pour estimer les caractéristiques de l'ensemble de la population.

F

Fichier analytique
Ensemble de **microdonnées** de Statistique Canada pour une enquête donnée, qui peut être utilisé dans les **centres de données de recherche** (CDR) au Canada. Les CDR permettent aux chercheurs d'accéder aux microdonnées d'enquêtes sur les ménages et sur la population. Les centres comptent des employés de Statistique Canada. Ils sont opérés en vertu des dispositions de la ***Loi sur la statistique*** et sont administrés conformément à toutes les règles de confidentialité. Ils ne sont accessibles qu'aux chercheurs dont les projets ont été approuvés et qui ont été assermentés comme « personnes réputées être employées » en vertu de la Loi sur la statistique.

Fichier de microdonnées à grande diffusion (FMGD)
Fichiers qui donnent accès aux **unités répondantes**, ce qui fait que les utilisateurs peuvent mener leur propre recherche ou analyse. Ils comprennent un ensemble de données non identifiables incluant les caractéristiques relatives aux unités de l'enquête (par exemple, personnes, ménages ou entreprises). La diffusion au public de ces ensembles de données a été autorisée par le Comité de la diffusion des microdonnées de Statistique Canada. Les ensembles de données ne comprennent aucun renseignement

confidentiel, étant donné que les identificateurs individuels ont été supprimés, et que toute combinaison de données ou de niveaux géographiques, qui pourrait potentiellement révéler l'identité d'une unité répondante, a été modifiée.

Fréquence
Nombre de fois qu'un événement ou qu'un élément se retrouve dans un **ensemble de données**.

G - H - I

Guides de l'utilisateur
Ces guides accompagnent les ensembles de données d'enquête de Statistique Canada, comme les fichiers d'analyse et les **fichiers de microdonnées à grande diffusion (FMGD)**, et fournissent des renseignements techniques détaillés nécessaires pour utiliser les données de façon appropriée. Le guide comprend généralement des renseignements importants à connaître avant l'analyse des données : variables de pondération utilisées, procédures relatives à l'estimation de la variance et précautions à prendre pour la diffusion des données.

Imputation
Remplacement de données manquantes ou invalides par des données valides. Elle est généralement exécutée selon des règles déterminées au préalable ou à partir de données d'un « voisin statistique », c'est-à-dire une autre **unité répondante** qui comporte des caractéristiques similaires. L'imputation est souvent combinée au **contrôle**.

Information
Données qui ont été enregistrées, classées, organisées, mises en correspondance ou interprétées à l'intérieur d'un cadre, afin d'obtenir une signification.

Institution
Les logements collectifs institutionnels comprennent les hôpitaux généraux et spécialisés, les hôpitaux pour personnes souffrant de maladies chroniques et hôpitaux de soins de longue durée, les établissements de soins infirmiers, les foyers collectifs ou établissements pour personnes avec incapacités physiques et centres de traitement, les foyers collectifs pour enfants et adolescents, les foyers collectifs ou institutions pour personnes ayant une déficience psychiatrique ou développementale, les établissements correctionnels fédéraux, les établissements de détention provinciaux et territoriaux, les établissements pour jeunes contrevenants, les prisons et lieux de détention de la police, les refuges pour personnes sans adresse fixe, les refuges pour femmes violentées et leurs enfants ainsi que les autres refuges offrant des services d'hébergement et d'aide.

Intelligibilité
Facilité avec laquelle l'utilisateur peut comprendre, utiliser de façon appropriée et analyser les données ou l'information. Le degré d'intelligibilité est déterminé dans une large mesure par : la pertinence des définitions des concepts, des populations cibles et des variables; la terminologie qui sous-tend les données, et les renseignements concernant les limites des données.

J - K - L

Loi sur la statistique
Loi concernant la statistique au Canada. Comprend la définition du mandat de Statistique Canada : « Est maintenu, sous l'autorité du ministre, un bureau de la statistique appelé Statistique Canada, dont les fonctions sont les suivantes :

- recueillir, compiler, analyser, dépouiller et publier des renseignements statistiques sur les activités commerciales, industrielles, financières, sociales, économiques et générales de la population et sur l'état de celle-ci;
- collaborer avec les ministères à la collecte, à la compilation et à la publication de renseignements statistiques, y compris les statistiques qui découlent des activités de ces ministères;
- recenser la population du Canada et faire le recensement agricole du Canada de la manière prévue à la présente loi;
- veiller à prévenir le double emploi dans la collecte des renseignements par les ministères;
- en général, favoriser et mettre au point des statistiques sociales et économiques intégrées concernant l'ensemble du Canada et chacune des provinces, et coordonner des projets pour l'intégration de telles statistiques. »

M

Marge d'erreur
Dans une enquête sur échantillon, les résultats de l'échantillon sont utilisés pour estimer ce que les résultats auraient été si l'ensemble de la population avait été mesuré. Dans ce processus d'**estimation**, un certain niveau d'erreur est inévitable. La marge d'erreur, utilisée pour former des intervalles de confiance, sert d'indicateur brut de la précision d'une estimation. Par exemple, les sondeurs disent souvent qu'un certain pourcentage de la population, plus ou moins la marge d'erreur (exprimée en points de pourcentage), devrait voter pour un certain candidat, 19 fois sur 20. Pour calculer la marge d'erreur, qui dans cet exemple correspond à un intervalle de confiance de 95 %, le sondeur utiliserait l'équivalent de plus ou moins deux erreurs types de l'estimation (voir **erreur type**).

Méthode bootstrap
Méthode servant à estimer l'erreur d'un ensemble de données associé à l'**échantillonnage**. L'échantillonnage entraîne des erreurs, parce que l'on ne recueille pas des données pour l'ensemble de la population, mais uniquement pour une sous-section, appelée échantillon, qui sert ensuite à produire des estimations pour l'ensemble de la population. Il existe plusieurs méthodes pour estimer le niveau d'erreur d'échantillonnage. Dans le cadre de la méthode bootstrap, on sélectionne habituellement un certain nombre de sous-échantillons de l'échantillon principal et on produit des estimations pour chaque sous-échantillon. L'erreur d'échantillonnage est estimée comme fonction des différences observées entre les estimations des différents sous-échantillons et de l'échantillon complet.

Méthodologie
Ensemble de méthodes et de techniques de recherche appliquées à un domaine particulier d'étude. À Statistique Canada, par méthodologie, on entend la méthodologie d'enquête.

Microdonnées
Fichier d'**enregistrements** se rapportant à des unités **répondantes** individuelles.

N – O

Observation
Données recueillies pour une variable donnée concernant une **unité répondante** particulière. Parmi les exemples figurent les valeurs particulières d'une unité répondante pour des caractéristiques comme l'âge, le sexe ou l'état matrimonial : l'observation pourrait être « 77 », « femme » et « veuve ».

P

Plan d'échantillonnage
Ensemble de spécifications qui décrivent les éléments d'échantillonnage d'une enquête de façon détaillée. Ces éléments comprennent la **population**, la **base de sondage**, les **unités** enquêtées par l'enquête, la taille de l'échantillon, la sélection de l'échantillon et la méthode d'**estimation**.

Poids
Nombre moyen d'unités de la population que l'unité de l'enquête représente. Parmi les exemples d'unités figurent une personne ou un ménage. Les poids sont appliqués aux **unités répondantes** dans un ensemble de données d'échantillon, afin de veiller à ce que, lorsque l'on fait des inférences à partir des données d'enquête en fonction de paramètres de population, les estimations des caractéristiques correspondent à la population totale.

Population
Groupe complet d'unités auxquelles les résultats de l'enquête s'appliquent. Ces unités peuvent être des personnes, des ménages, des entreprises, des institutions, etc. Le terme « population cible » est souvent utilisé pour désigner toutes les **unités** pouvant être enquêtées, conformément à la définition claire et précise de l'étude de l'enquête. Il s'agit de la population au sujet de laquelle on veut obtenir des données.

Produit d'information
Organisation des résultats des activités de Statistique Canada, y compris les fichiers de données, les **bases de données**, les tableaux, les graphiques, les cartes et le texte. Cette organisation peut être définie au préalable (produit d'information standard) ou faire suite à des demandes spéciales (produit d'information personnalisé). Les produits d'information sont disponibles sous forme imprimée ou électronique.

Proportion
Relative au nombre de réponses qui se retrouvent dans une catégorie de réponse donnée, en rapport avec les réponses totales. Elle est calculée en divisant la fréquence de la catégorie de réponse par le nombre total de réponses à la question.

Q – R

Qualité des données
Degré ou niveau de confiance quant à l'adaptation des données et des renseignements statistiques à leur utilisation. Les questions particulières de qualité ou d'adaptation à l'utilisation qui doivent être prises en compte par Statistique Canada sont la pertinence, l'exactitude, l'actualité, l'accessibilité, l'**intelligibilité** et la cohérence.

Recensement de la population
Collecte de données concernant toutes les unités d'une population, et est aussi parfois appelé enquête auprès d'un échantillon de 100 %. En vertu de la ***Loi sur la statistique*** de 1971, il est obligatoire de mener un recensement à l'échelle du pays tous les cinq ans. Le recensement de la population fournit des

données qui servent à planifier des services communautaires comme des écoles, garderies, services de police et de protection contre les incendies ; à prévoir la demande des consommateurs ; et à procéder à des études de marché.

Région métropolitaine de recensement (RMR) et agglomération de recensement (AR)
Territoire formé d'une ou de plusieurs municipalités voisines les unes des autres qui sont situées autour d'un grand noyau urbain. Une région métropolitaine de recensement doit avoir une population d'au moins 100 000 habitants et le noyau urbain doit compter au moins 50 000 habitants. L'agglomération de recensement doit avoir un noyau urbain d'au moins 10 000 habitants.

Régression
Méthode statistique servant à prédire la valeur d'une caractéristique, grâce à l'étude de son rapport avec une ou plusieurs autres caractéristiques. Ce rapport est exprimé grâce à une équation de régression.

Régression logistique
Il s'agit d'une forme d'analyse de **régression** utilisée lorsque la variable de réponse est une **variable** binaire (une variable comportant deux valeurs possibles).

Répondant
Personne qui fournit les données pour l'**unité enquêtée**, qu'il s'agisse d'une personne, d'un ménage, d'une entreprise ou d'une institution.

S

Suppression
Processus grâce auquel des données particulières ne sont pas diffusées, sur la base de critères conçus pour protéger la confidentialité. La suppression de « cellules » se rapporte aux procédures utilisées pour protéger les données de tableaux de nature délicate de la divulgation, une cellule étant une entrée dans un tableau. Dans le cas de l'ECI, les données ont aussi été supprimées pour des raisons de qualité (C.V. plus élevé que 33,3 %).

T - U – V

Taux d'échantillonnage ou fraction de sondage
Taille d'échantillon divisé par la taille de la population.

Taux de réponse
Proportion d'un échantillon pour laquelle une réponse au questionnaire est obtenue, habituellement exprimée en pourcentage. La non-réponse englobe ceux qui refusent de participer, ainsi que les personnes que l'on n'a pas pu contacter dans le cadre de l'enquête.

Unité
Voir unité enquêtée.

Unité d'échantillonnage ou unité échantillonnée
Unité sélectionnée au moyen du **plan d'échantillonnage** et au sujet de laquelle des mesures sont recueillies dans le cadre d'une enquête. Parmi les exemples figurent les personnes, les ménages, les familles ou les entreprises. Dans le cas de l'ECI, l'unité d'échantillonnage est la personne.

Unité enquêtée
Unité sélectionnée pour laquelle des mesures sont recueillies pour une enquête sur échantillon ou un recensement. Parmi les exemples figurent les personnes, les ménages, les familles ou les entreprises. Dans le cas de l'ECI, l'unité enquêtée (qui est aussi l'unité échantillonnée, étant donné que l'ECI est une enquête sur échantillon) est la personne agée de 15 ans et plus.

Variable
Caractéristique qui peut prendre plus d'une valeur, à laquelle une mesure numérique est attribuée (par exemple, revenu, âge et poids).

Variable dérivée
Nouvelle **variable** élaborée grâce à l'application d'opérations logiques ou mathématiques à une ou plusieurs variables existantes, afin de répondre à un besoin de données particulier. Par exemple, une variable de l'âge peut être calculée à partir des données sur la date de naissance. Par ailleurs, une variable dérivée appelée

« présence d'un problème de santé chronique » peut être obtenue selon qu'un répondant a répondu « oui » au moins une fois à une série de questions concernant des problèmes de santé chronique particuliers, comme l'asthme, le diabète, la maladie cardiaque, etc.

Variance
Mesure de dispersion d'une caractéristique ou d'une variable donnée dans un ensemble de données. Elle montre dans quelle mesure une variabilité existe pour cette caractéristique. Au niveau technique, elle est calculée comme étant la moyenne des écarts au carré entre chaque observation et leur moyenne sur l'ensemble de données pour une variable particulière.

Références

BERNIER, J., et K. NOBREGA. 1998. *Outlier detection in asymmetric samples: A comparison of an inter-quartile range method and a variation of a sigma-gap method*, congrès annuel de la Société statistique du Canada, juin 1998.

HADDOU, M. 2013. *Bootstrap Variance Estimation Specifications – Aboriginal Peoples Survey*, document interne, janvier 2013.

LANGLET, É., J.-F. Beaumont et P. Lavallée. 2008. *Bootstrap Methods for Two-Phase Sampling Applicable to Postcensal Surveys*, article présenté au Comité consultatif des méthodes statistiques, mai 2008, Ottawa.

MACKENZIE, A., M. HURST et S. CROMPTON. 2009. *Définition de l'incapacité dans l'Enquête sur la participation et les limitations d'activités, Tendances sociales canadiennes*, nº 88, « Série sur la vie avec une incapacité », produit nº 11-008-X au catalogue de Statistique Canada, décembre 2009.

MOREL, J., et C. NAMBEU. 2013. *National Household Survey: Weighting and estimation update*, présenté au congrès de la Société statistique du Canada en mai 2013.

RESSOURCES HUMAINES ET DÉVELOPPEMENT DES COMPÉTENCES CANADA. 2010. Rapport fédéral sur les personnes handicapées : Rapport annuel du gouvernement sur la condition des personnes handicapées – 2010, Ressources humaines et Développement des compétences Canada, Gatineau, Québec.